Methods of Computational Neuroscience

Part 6 of 6

Marine Biological Laboratory - Woods Hole, MA
August 2009

Day 3 (August 19)

Bence Olveczky

- woods hole09.pdf

Ila Fiete

- 2009_aug_MBL_birdsong_topost.pdf

Day 4 (August 20)

Reza Shadmehr

- Woods_hole_09_2.ppt

Emo Todorov

- MBL09.pdf

Day 5 (August 21)

Simon Laughlin

- mcn 09 simon.pdf

Week 4 (Aug. 24-28)

Day 1 (Aug. 24)

Mitya Chklovskii

- Theoretical Connectomics.pdf

Day 2 (August 18)

Sara Solla

Haim Sompolinsky:

- Learning spike based rules-THE TEMPOTRON
- guetig_sompolinsky06_full.pdf
- gutig_sompolinsky_plos09.pdf

Larry Abbott's second talk: Chaotic networks

- AbbottTalk2.pdf
    - Jaeger H., and Haas H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science 304, 78-80.

Problem Sets

Week 1

Simulated Neurons

Problem Set 1.zip

See the papers from day 1 (Fairhall et al., 2006 and Pillow et al., 2008)

Information theory

ProblemSet-InformationTheory.pdf

ProblemSet-FisherInfoCorrelatedNeurons-MCN2009.pdf (updated 2009.08.11)

Solutions

ProblemSetSolution-InformationTheory-MCN2009.pdf (updated 2009.08.11)

ProblemSetSolution-FisherInfoCorrelatedNeurons-MCN2009.pdf (updated 2009.08.11)

Spectral analysis

Link to Chronux download website

Kleinfeld_Mitra_CSHL_2009_revised..pdf

matlab_for mcn_2990.m

example_data.mat

whisking.mat

090606_PA5_noStim.mat

Week 2

Working with XPP

NOTE: These problems are to be done at your discretion, and will not be presented.

- tcurrent.ode Exploring T-current and bifurcations
- ping.ode A simple model for PING

These are text files. The instructions are at each individual file after the 'Done' statement.

NEURON tutorial and exercises

- NEURON Tutorial ppt.pdf
- NEURON short exercises.pdf
  - Complete these exercises to become familiar with NEURON, as well as gaining a better understanding of dendritic dynamics.
- simple_layerVpyramidal.zip

- Webpage for downloading NEURON to your computer
  - Follow the instructions for your OS of choice. At the computer stations in the lab we use the MacOS-X 10.4 version
  - For Mac laptops, you HAVE to install the XCode package (size is about 1GB). Please ask the TA's for it, before the tutorial on Wednesday.

NEURON mini projects

- NEURON mini projects.pdf
- Rall_MCN_Assignment.pdf
- Katz et al 2009 - Neuron.pdf
- Kampa & Stuart 2006 - JNeurosci.pdf

This version of the code contains an additional .hoc file for use with the second mini project. Open "exerciseS2.hoc" to access it.

mbl0813.zip

for problem 3:
Calcium spikes in basal dendrites (Kampa and Stuart 2006) [modelDB]

Attractor network exercises

Mark Goldman will discuss your solutions with you on Monday afternoon.

MCN_RecurrentNetsExercise.pdf

MCN 2009 Lecture Notes & Papers
Reader – Part 1 of 6

Week 1 (Aug. 3-7)

Day 1 (Aug. 3) - Reader – Part 1 of 6

Adrienne Fairhall

- mcn09FairhallSlides.pdf
- Fairhall2006.pdf (Covariance analysis in retina)
- Pillow2008.pdf (GLM analysis in retina)
- Fairhall-etal-01.pdf (Multiple timescale adaptation in the fly)
- lundstromNN08.pdf (Fractional differentiation in cortical neurons)
- MaravallPLOS07.pdf (Gain control in barrel cortex)
- ruderman1997.pdf (Review of natural scene statistics)

Bill Bialek (website)

- bialek_02.pdf Les Houches: *Thinking about the brain*
- The ideas about optimal motion estimation go back to a paper with Marc Potters, which can be found here: http://www.princeton.edu/~wbialek/our_papers/potters+bialek_94.pdf
- The material on photon counting in vision is covered in a set of lecture notes which can be found at

  http://www.princeton.edu/~wbialek/PHY562.html

Day 2 (Aug. 4) - Reader – Part 2 of 6

Jonathan Pillow

- pillow_slides09_MaxLi_GLM.pdf (lecture slides)

Michael Berry (website)

- MCN 2009 Populations.ppt
- Recording All Cells.pdf
- Redundancy in Retinal Code.pdf

Rava da Silveira

- TheOrchestralBrain-WithAnnotations.pdf

Day 3 (Aug. 5)

Roozbeh Kiani

- Kiani_CategorizationInIT.pdf
- kianiJNP07.pdf
- kianiNeuron08.pdf

Jim DiCarlo

- DiCarloWHslides.pdf

Elad Schneidman

- Schneidman-MBL-MCN-2009.pdf
- Schneidman_2006_Nature_Weak pairwise correlations imply strongly.pdf
- Schneidman_PRL_2003_Network information and connected correlations.pdf
- Tkacik+al_06.pdf

Day 4 (Aug. 6)

Ehud Ahissar

- Ahissar_MCN2009_partA.pdf
- Ahissar_MCN2009_partB.pdf
- Knutsen & Ahissar TINS 2009 + Erratum.pdf
- Diamond et al NRN 2008.pdf

David Kleinfeld (website)

- MCN_2009_1.pdf
- MCN_2009_2_online.pdf
- john_curtis_NN_touch_S1 .pdf
- david_Encyclopedia_of_Neuroscience.pdf
- david_ehud_mat_Current_Opinions_06.pdf

Day 5 (Aug 7)

Bill Spain

- MCN 09short.ppt
- Reading list for Bill Spain's seminar.doc
- Jeffress.pdf (one of the papers on the reading list that is not available online)

Week 2 (Aug. 10-14)

Day 1 (Aug 10)

Bard Ermentrout

- chapt1.pdf
- A very comprehensive tutorial on bifurcations for neuro
- A dated tutrial for XPP
- A very quick look at using xpp to solve the HH eqns
- A link to the xpp homepage

ml.ode Morris-Lecar ODE
hhh.ode Hodgkin-Huxley

Installation of XPP on your computer:

- Using XPP on windows

Adrienne Fairhall

- mcn2009Fairhall_2.pdf
- hongAgueraFairhallNC07.pdf
- lundstromFamulareFairhall.pdf
- agueraFairhallBialek.pdf
- hongLundstromFairhallPloSCB08.pdf
- lundstromNC08.pdf

Day 2 (Aug 11)

Dan Johnston

Day 3

Michael Hausser

- Hausser_slides.pdf
- Sjostrom_Hausser_Neuron06.pdf
- Hausser&London2005review.pdf

- mageeNatRevNeuro00.pdf

Bartlett Mel

- Mel_MCN_lecture.pdf

Day 4 (August 13)

Emre Aksay

- MCNTalk_2009_Aksay_Wiki.pptx

Mark Goldman

- Goldman_MCNTalk_Theory09.pdf

Day 5 (August 14)

Nancy Kopell

43:01

53:09

Sara A. Solla

123:21

- MBL09_SOLLA.ppt
- PRL04.pdf
- Chaos07.pdf

Week 3 (Aug. 17-21)

Day 1 (Aug. 17)

Karel Svoboda

- Svoboda_MCN.pdf

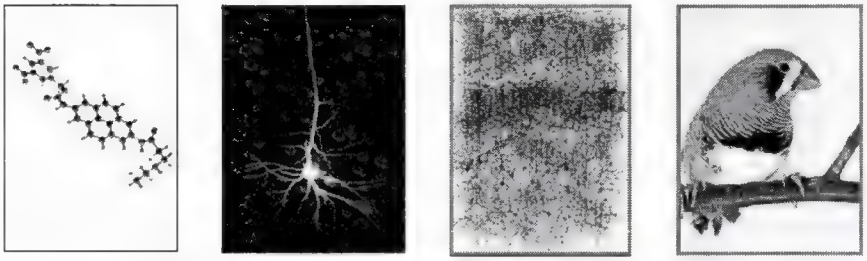
Larry Abbott

- AbbottLect1.pdf
- force5.pdf
- Bi, G.-q. & Poo, M.M. Synaptic modification by correlated activity: Hebb's postulate revisited. Annu. Rev. Neurosci. 24,139-166.

- Caporale N. & Dan, Y. Spike Timing-Dependent Plasticity: A Hebbian LearningRule. Annu. Rev. Neurosci. 31, 25-46 (2008).
- Dan Y. & Poo M.M. Spike timing-dependent plasticity: from synapse to perception. Physiol Rev. 86, 1033-1048 (2006).

Neural mechanisms underlying the learning and execution of learned vocalizations in the songbird

Molecule | Neuron | Circuit | Behavior



Bence Ölveczky
Department of Organismic and Evolutionary Biology
Center for Brain Science
Harvard University

Using the songbird as a model system

Birdsong is a complex learned motor behavior

Well defined and ethologically relevant behavior

The behavior is robustly expressed in the lab

The behavior is easily measured and quantified

Behavior is implemented by dedicated neural circuits

Brain regions are accessible for recording/manipulation during behavior

Using the songbird as a model system

Birdsong is a complex learned motor behavior

Well defined and ethologically relevant behavior

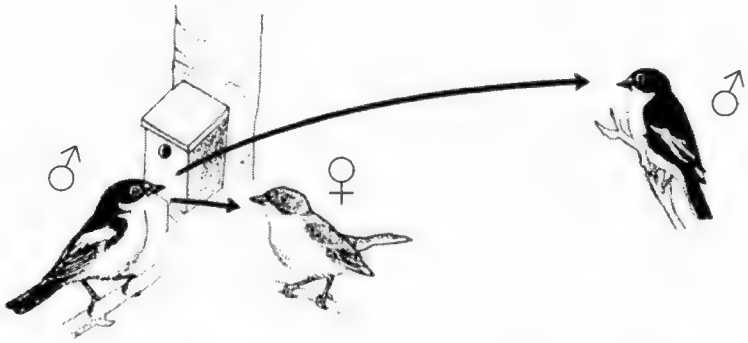
The behavior is robustly expressed in the lab

The behavior is easily measured and quantified

Behavior is implemented by dedicated neural circuits

Brain regions are accessible for recording/manipulation during behavior

Birdsong is a "dual function" signal

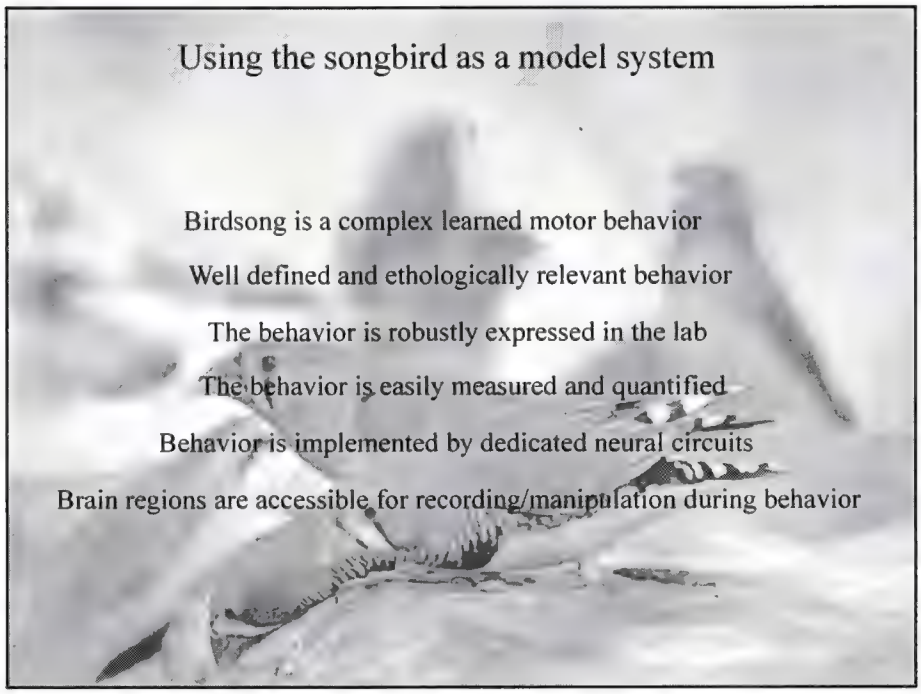


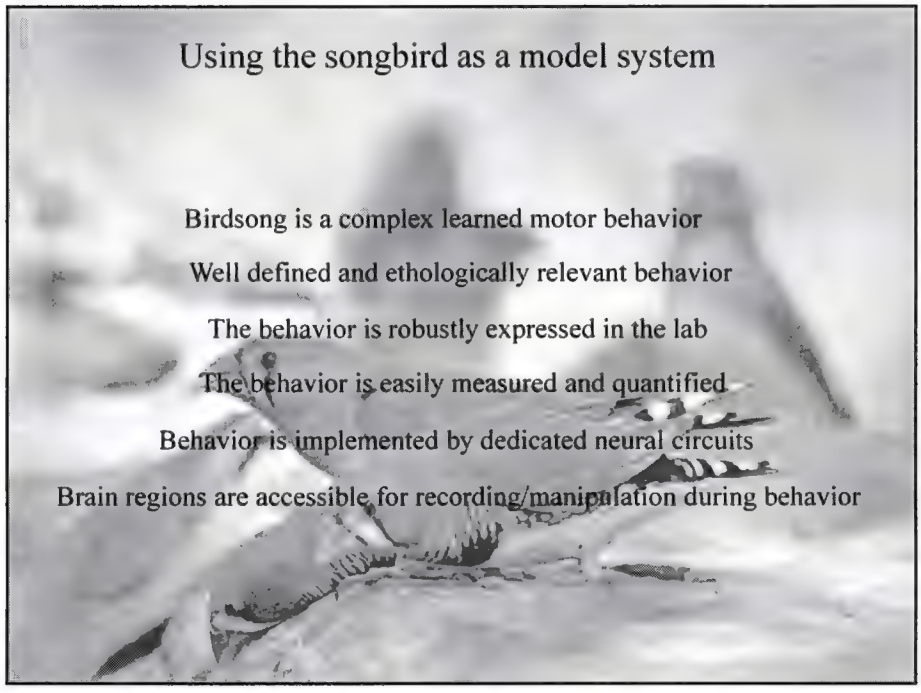
MALE → FEMALE

- attract mates
- stimulate courtship & reproduction

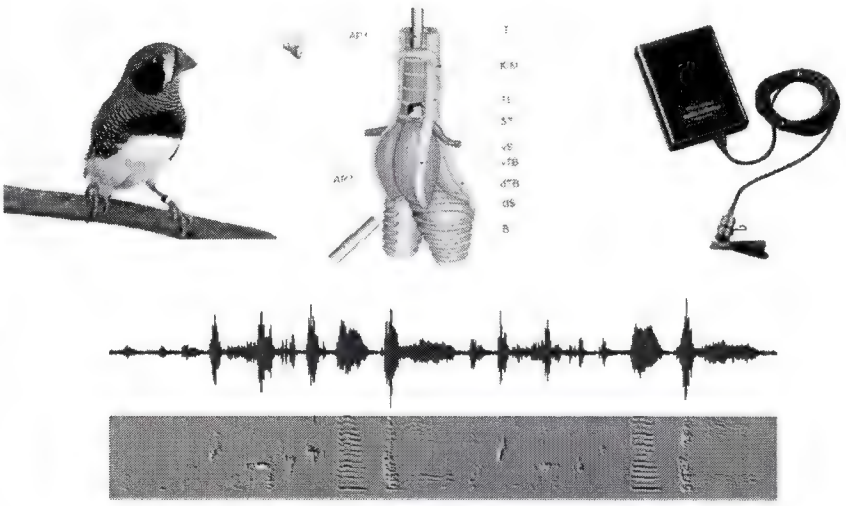
MALE → MALE

- repel rivals from defended space


Using the songbird as a model system
Birdsong is a complex learned motor behavior
Well defined and ethologically relevant behavior
The behavior is robustly expressed in the lab
The behavior is easily measured and quantified
Behavior is implemented by dedicated neural circuits
Brain regions are accessible for recording/manipulation during behavior


Using the songbird as a model system
Birdsong is a complex learned motor behavior
Well defined and ethologically relevant behavior
The behavior is robustly expressed in the lab
The behavior is easily measured and quantified
Behavior is implemented by dedicated neural circuits
Brain regions are accessible for recording/manipulation during behavior

A simple measure of a complex behavior



Using the songbird as a model system

Birdsong is a complex learned motor behavior

Well defined and ethologically relevant behavior

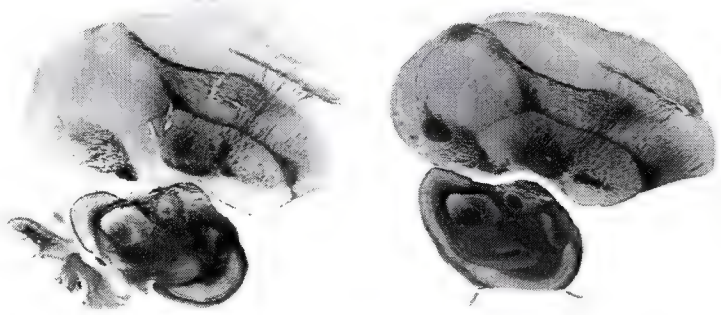
The behavior is robustly expressed in the lab

The behavior is easily measured and quantified

Behavior is implemented by dedicated neural circuits

Brain regions are accessible for recording/manipulation during behavior

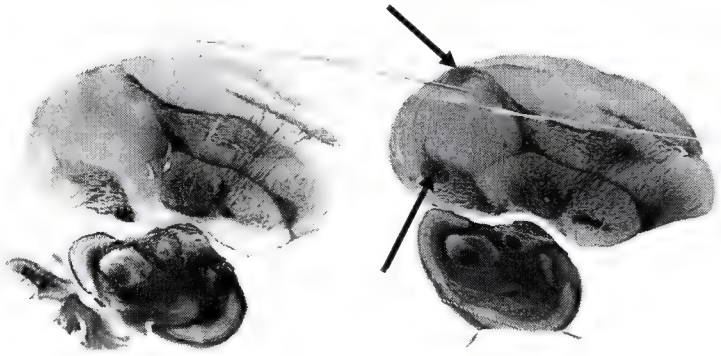
Neural circuitry underlying birdsong



female male

- Zebra finches are sexually dimorphic

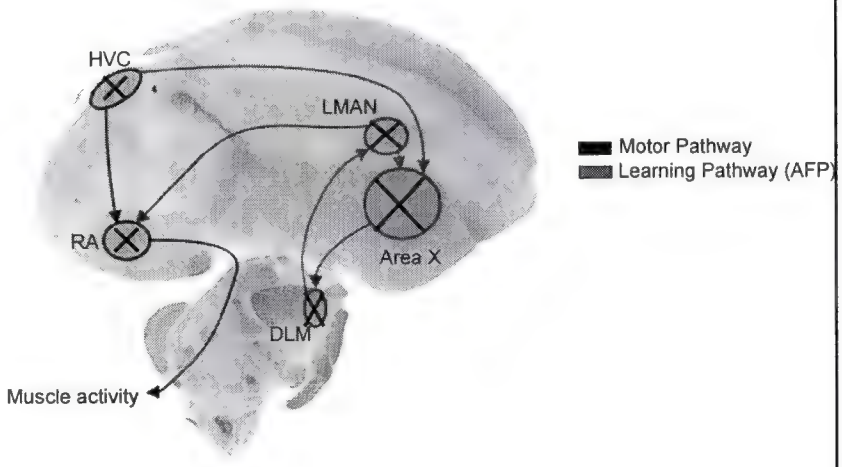
Neural circuitry underlying birdsong



female male

- Zebra finches are sexually dimorphic

Neural circuitry underlying birdsong



Using the songbird as a model system

Birdsong is a complex learned motor behavior

Well defined and ethologically relevant behavior

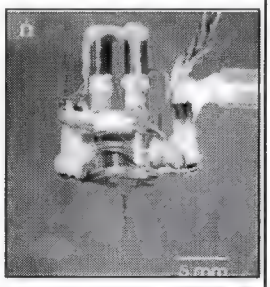
The behavior is robustly expressed in the lab

The behavior is easily measured and quantified

Behavior is implemented by dedicated neural circuits

Brain regions are accessible for recording/manipulation during behavior

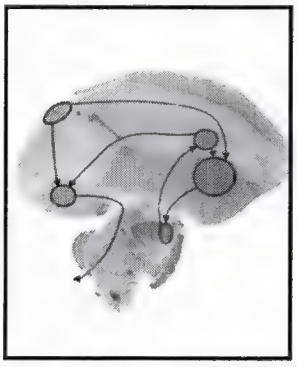
Recording the neural control signals during singing



Leonardo and Fee 2005

Reverse Engineering the Songbird

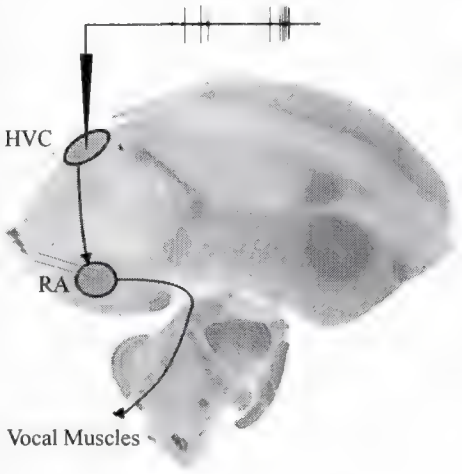
Neural Circuit

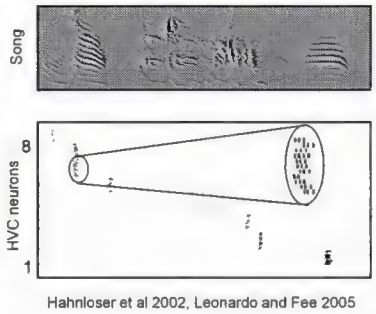


Behavior

Manipulate & Measure

Motor control of song in adult birds





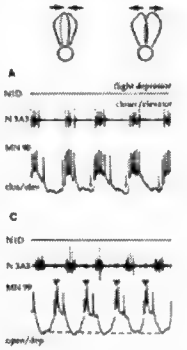
Hahnloser et al 2002, Leonardo and Fee 2005

HVC- no known map
RA- muscle map

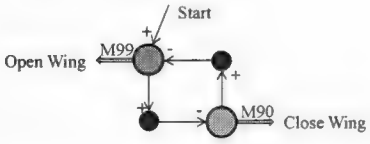
Synfire chain, Abeles, 1982

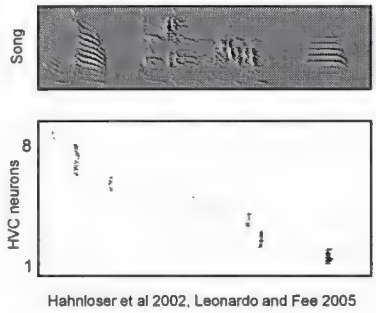
Is the temporal sequence generated within HVC?

Cricket Stridulation

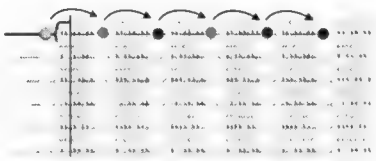


12° C
18° C
24° C
30° C



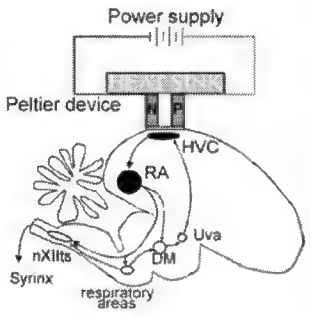


Hahnloser et al 2002, Leonardo and Fee 2005

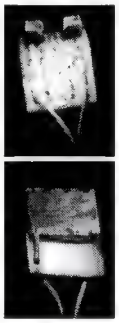


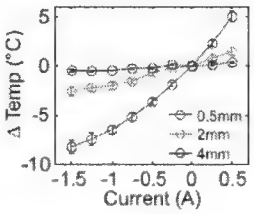
Synfire chain, Abeles, 1982

Cooling down HVC



Peltier Cooler



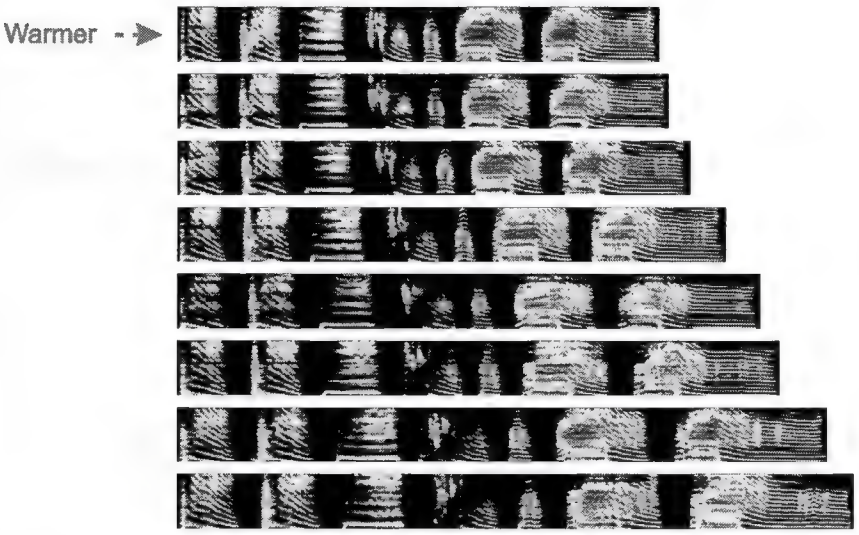


Long and Fee, 2008

Speed of propagation in HVC

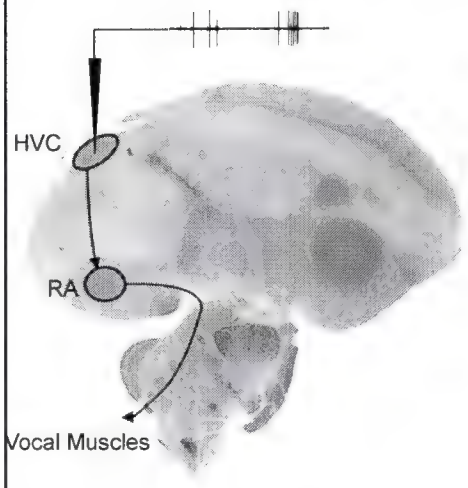
Normal Cooled

HVC generates the temporal sequence

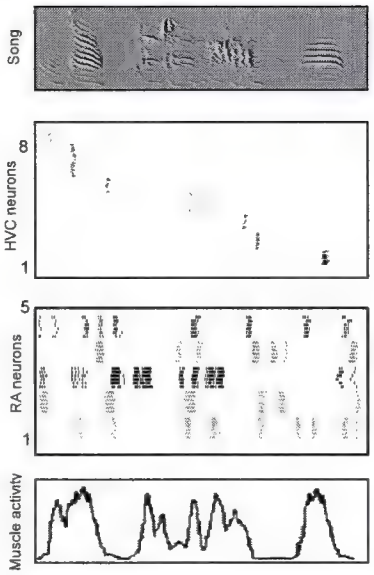


Cooling HVC (slowing synaptic transmission) slows down the song

Motor control of song in adult birds

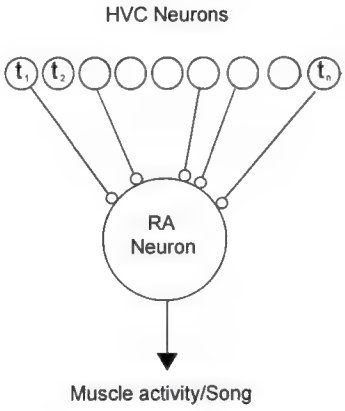


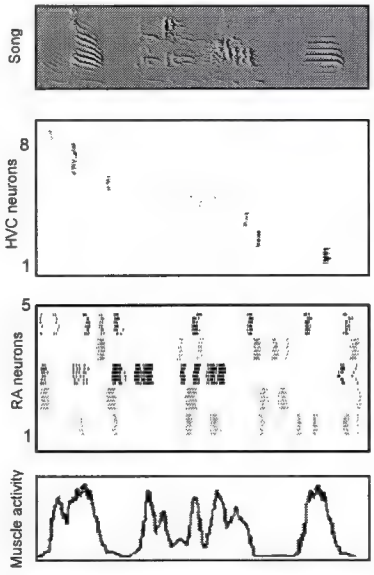
HVC- no known map
RA- muscle map



Hahnloser et al 2002, Leonardo and Fee 2005

Motor control of song in adult birds



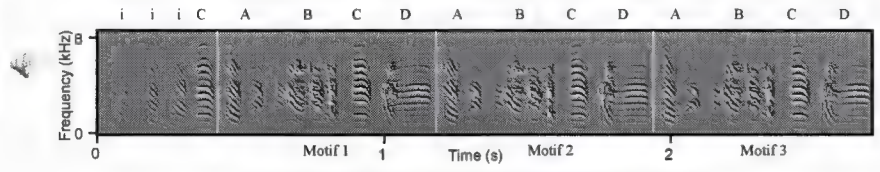


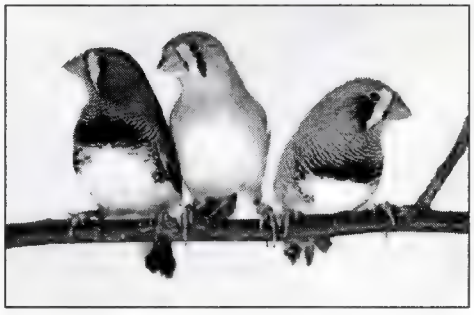
Hahnloser et al 2002, Leonardo and Fee 2005

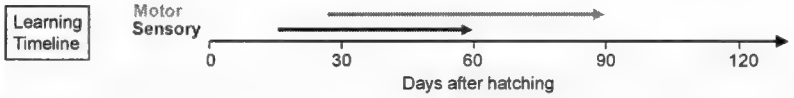
HVC – Conductor (temporal sequence)

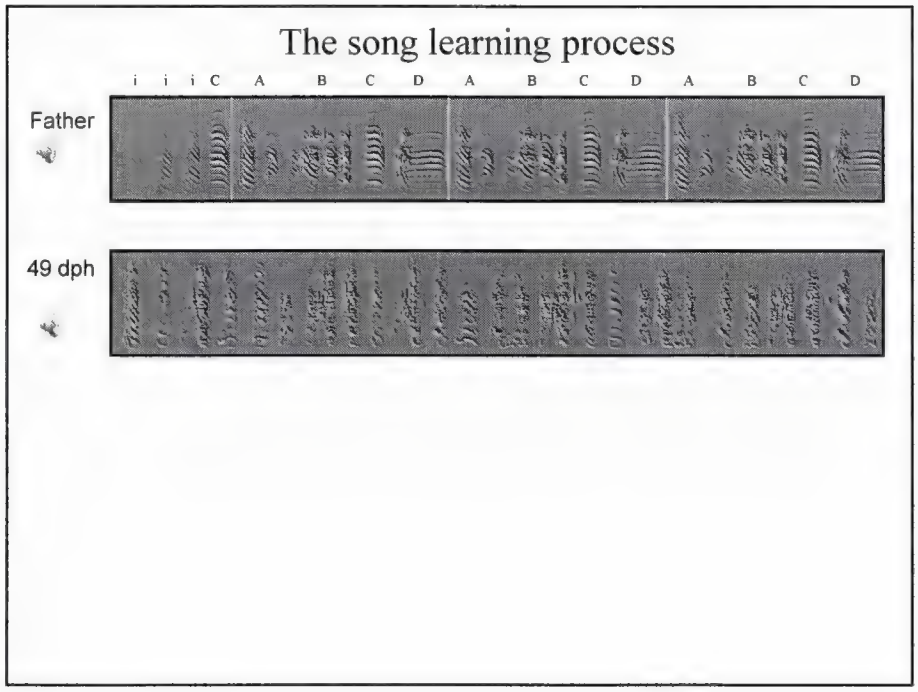
RA – Orchestra (translates the sequence into sound)

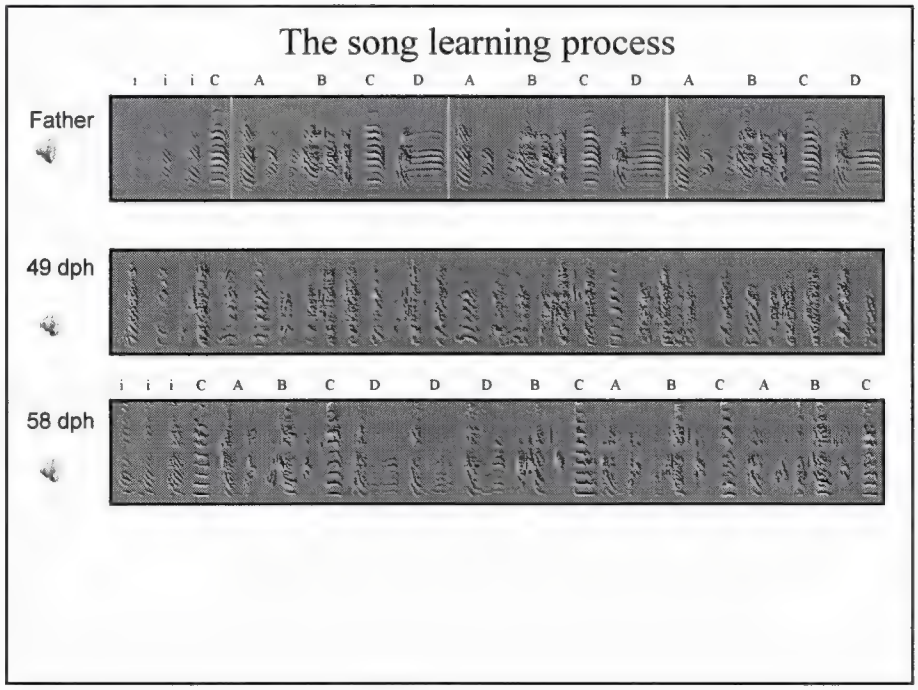
The song learning process



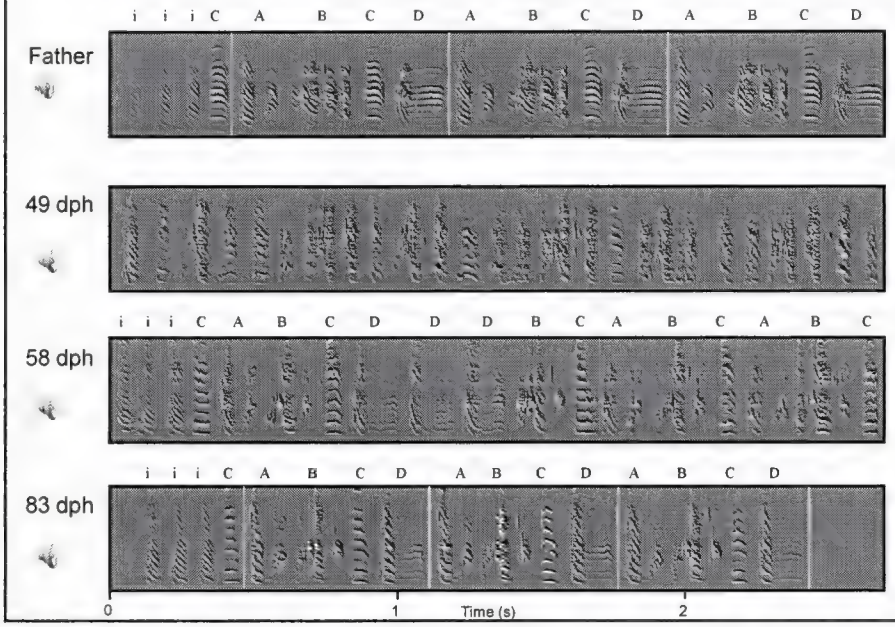





The song learning process
i i i C A B C D A B C D A B C D
Father
49 dph


The song learning process
i i i C A B C D A B C D A B C D
Father
49 dph
i i i C A B C D D D B C A B C A B C
58 dph

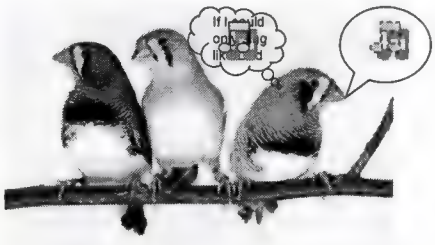
The song learning process



Reinforcement learning as a framework for song learning

Reinforcement learning (Sutton & Barto, 1998).

"In reinforcement learning the learner is not told which actions to take, but instead must discover which actions yield the most reward by trying them. These two characteristics - trial-and-error search and delayed reward - are the two most important distinguishing features of reinforcement learning."

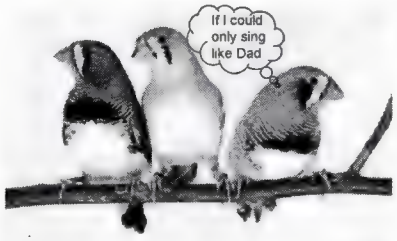


Explore motor programs → Get performance based feedback → Reinforce 'right" motor programs

Reinforcement learning as a framework for song learning

Reinforcement learning (Sutton & Barto, 1998).

"In reinforcement learning the learner is not told which actions to take, but instead must discover which actions yield the most reward by trying them. These two characteristics - trial-and-error search and delayed reward - are the two most important distinguishing features of reinforcement learning."

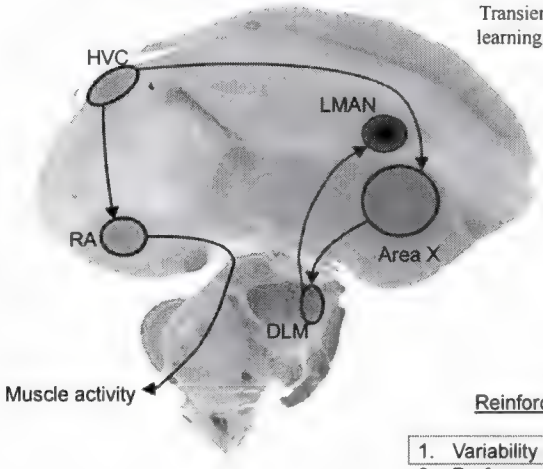


Explore motor programs → Get performance based feedback → Reinforce 'right" motor programs

What is the role of the AFP in song learning?

TTX or Muscimol

Transiently inactivating the output of the learning pathway in the juvenile songbird

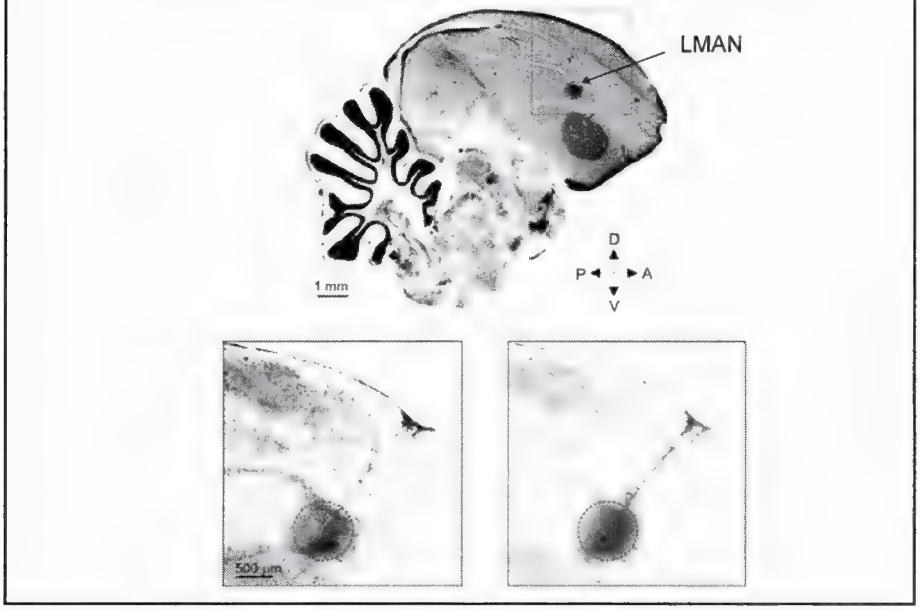


Motor Pathway
Learning Pathway (AFP)

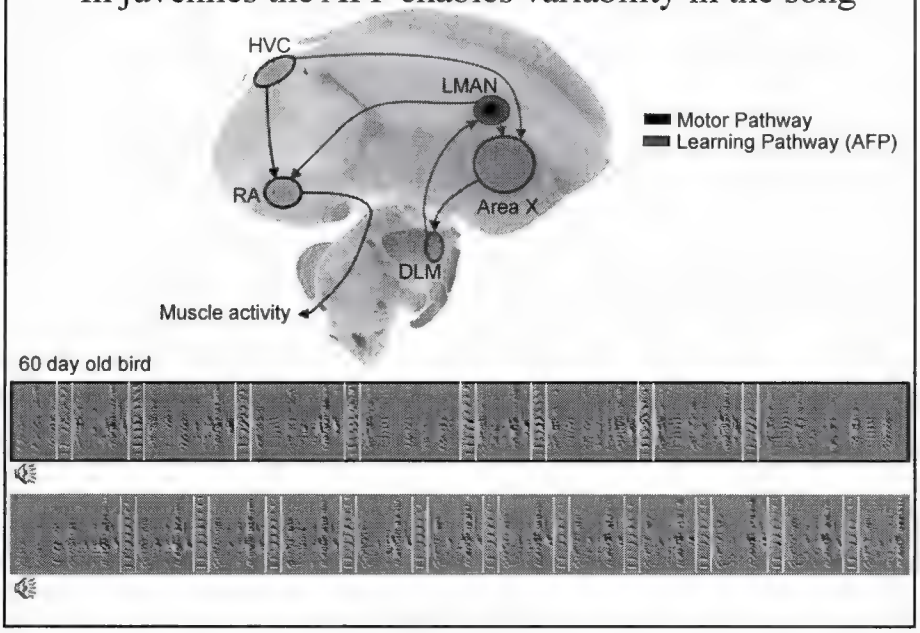
Reinforcement learning

1. Variability in the motor program
2. Performance based evaluation signal

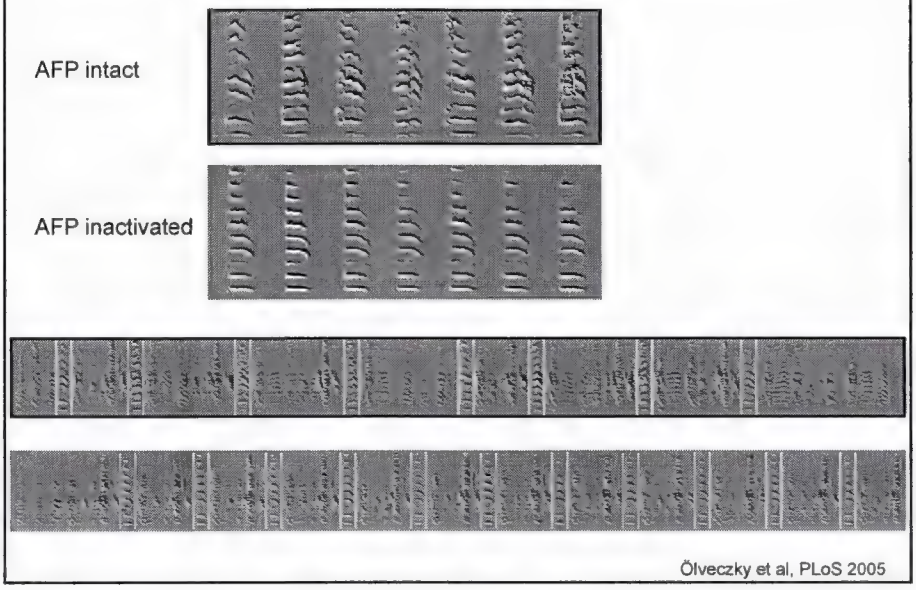
Transiently inactivating the learning pathway



In juveniles the AFP enables variability in the song

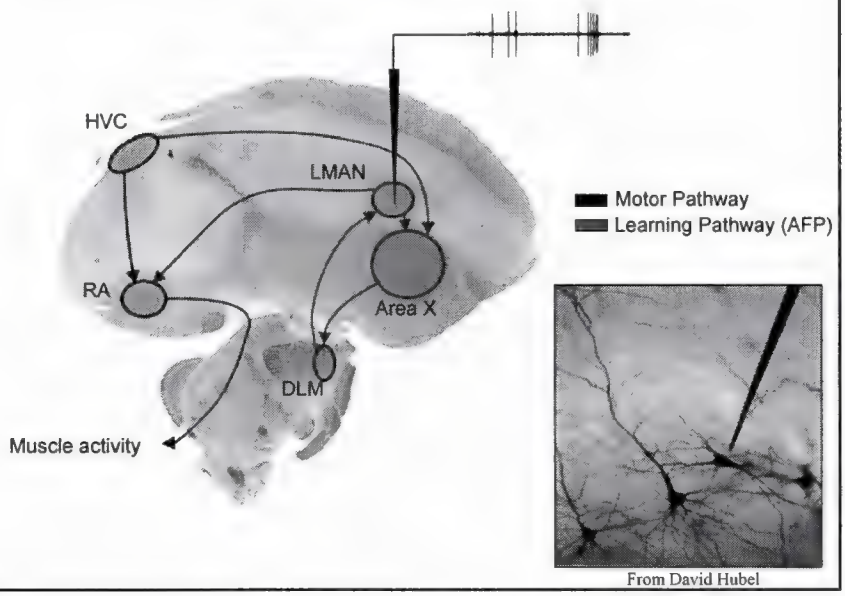


In juveniles the AFP enables variability in the song



Ölveczky et al, PLoS 2005

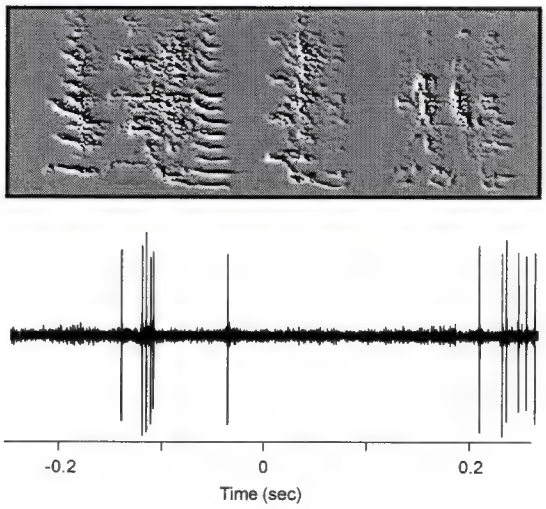
How is motor variability expressed in LMAN?



From David Hubel

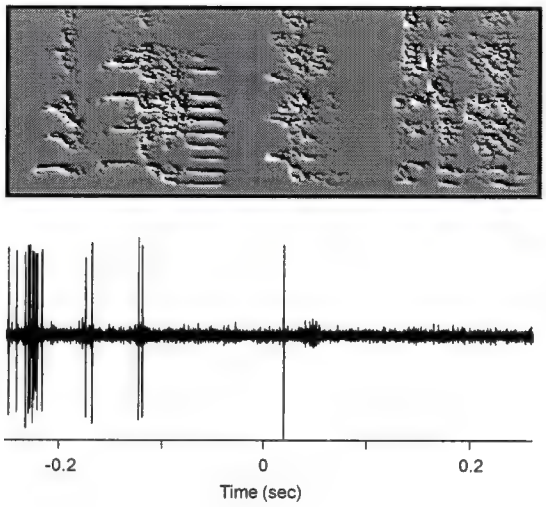
Recording from LMAN neurons in juvenile bird

Cell #1
Motif #1



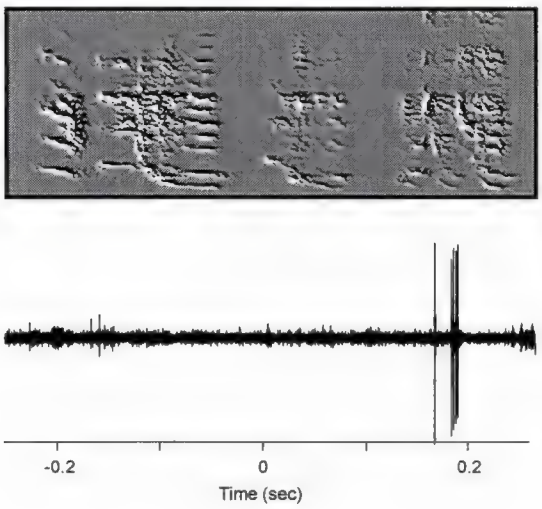
Recording from LMAN neurons in juvenile bird

Cell #1
Motif #2



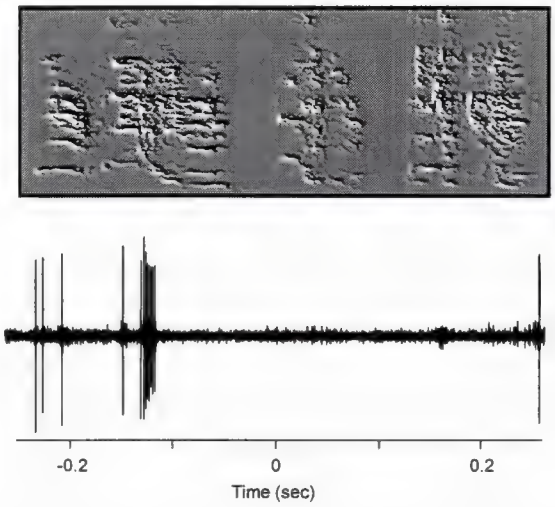
Recording from LMAN neurons in juvenile bird

Cell #1
Motif #3

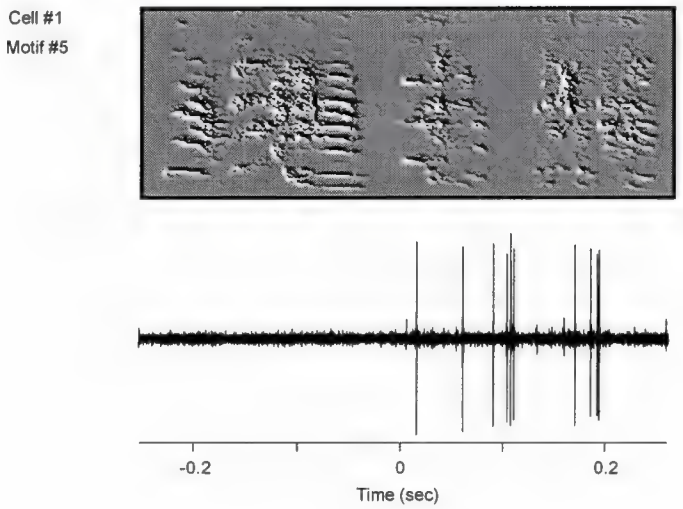


Recording from LMAN neurons in juvenile bird

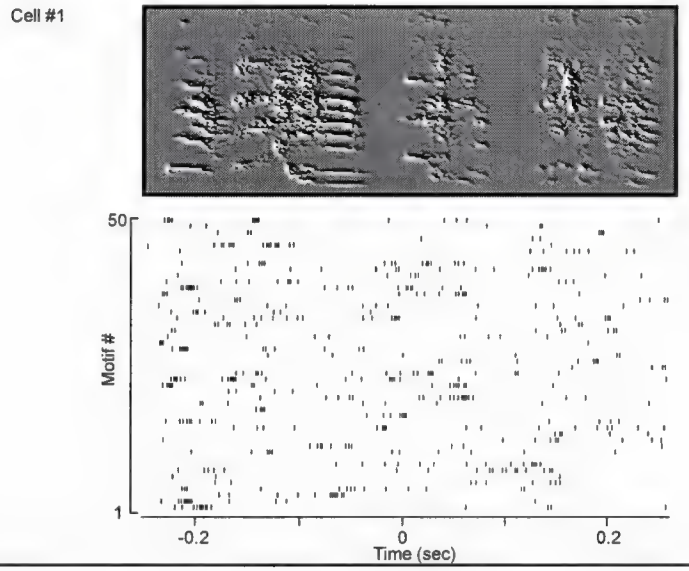
Cell #1
Motif #4



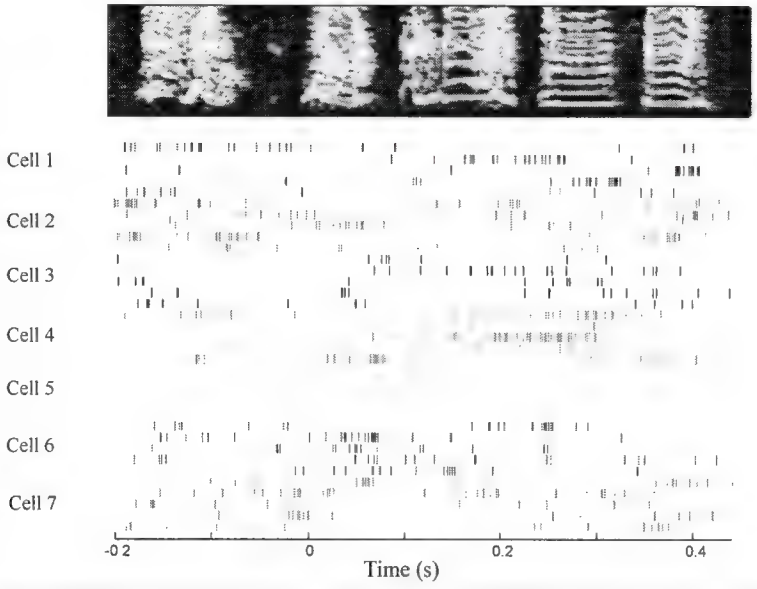
Recording from LMAN neurons in juvenile bird



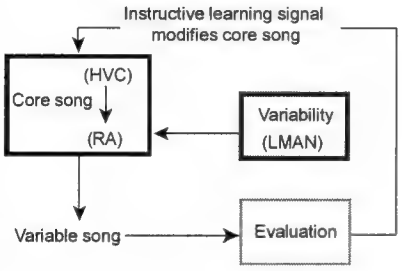
LMAN firing patterns are variable relative to song

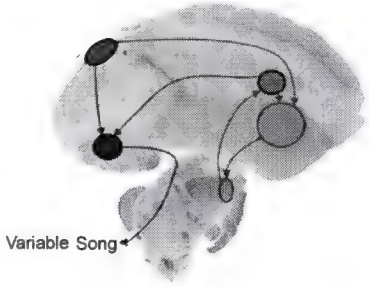


LMAN firing patterns are variable relative to song



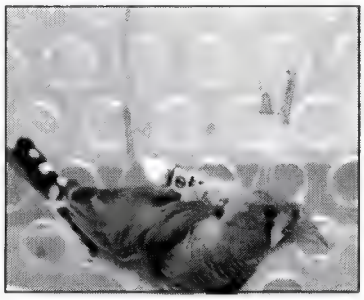
Reinforcement learning algorithm

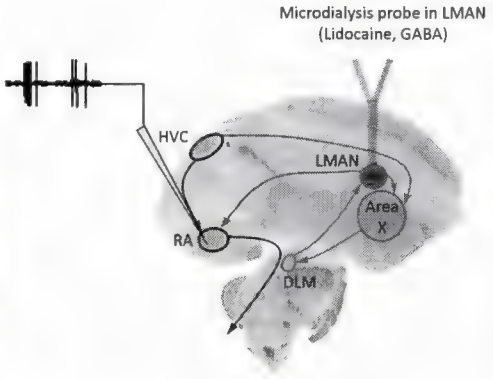




How does the variable input from LMAN influence the motor program in RA during learning?

How does LMAN input interact with the motor pathway during learning?



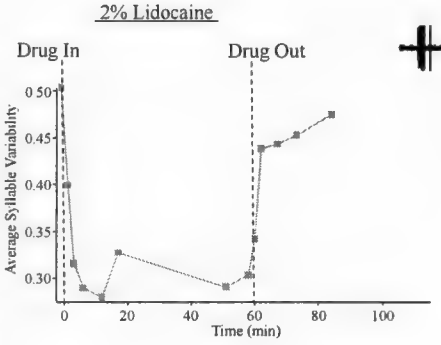


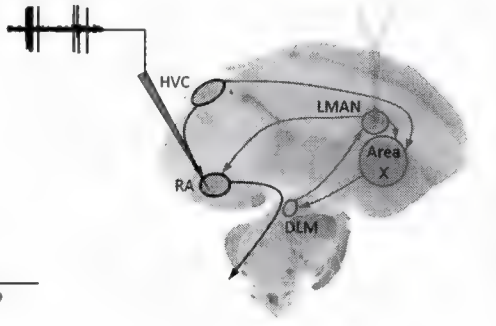
Record from RA neuron ——

Activity in LMAN ON

Song Variable

Recording from RA while turning LMAN off and on

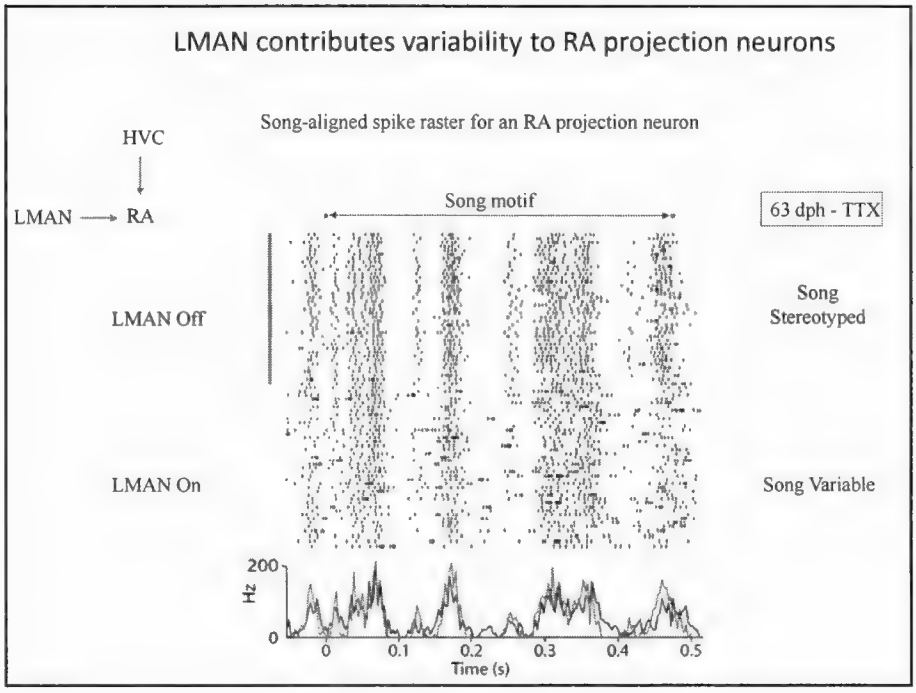


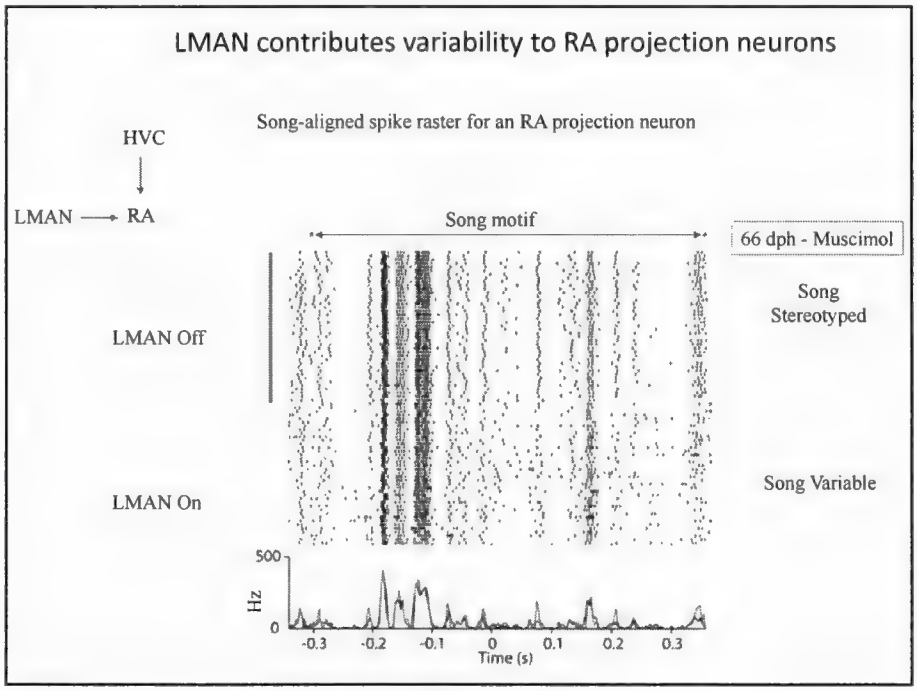


Record from RA neuron ——→

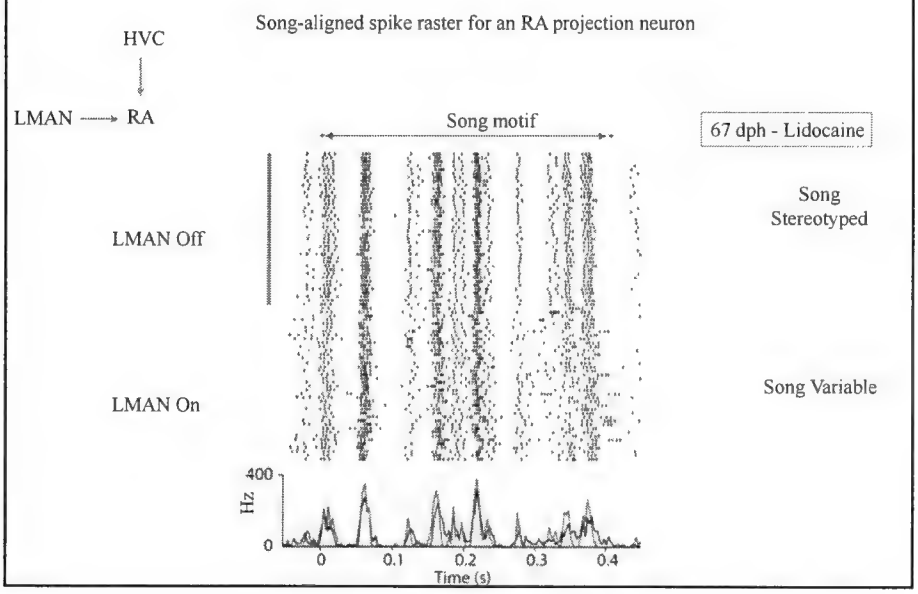
Activity in LMAN ON OFF ON

Song Variable Stereotyped Variable

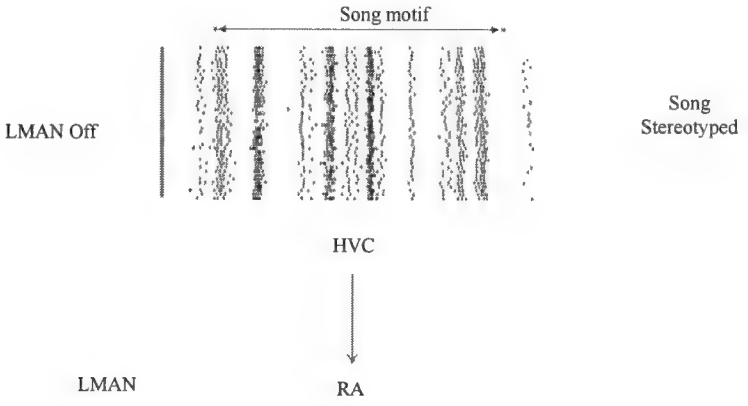

LMAN contributes variability to RA projection neurons
Song-aligned spike raster for an RA projection neuron
HVC
LMAN
RA
Song motif
63 dph - TTX
LMAN Off
Song Stereotyped
LMAN On
Song Variable
200
Hz
0
0
0.1
0.2
0.3
0.4
0.5
Time (s)


LMAN contributes variability to RA projection neurons
Song-aligned spike raster for an RA projection neuron
HVC
LMAN
RA
Song motif
66 dph - Muscimol
LMAN Off
Song Stereotyped
LMAN On
Song Variable
500
Hz
0
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
Time (s)

LMAN contributes variability to RA projection neurons



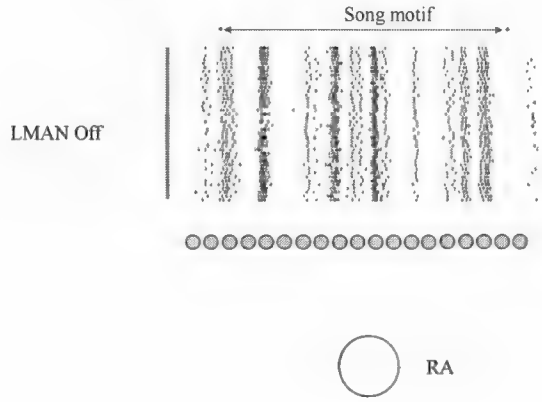
Shaping of the RA motor program in juvenile birds



Assumption: RA firing events are driven from HVC

Input from HVC is temporally precise

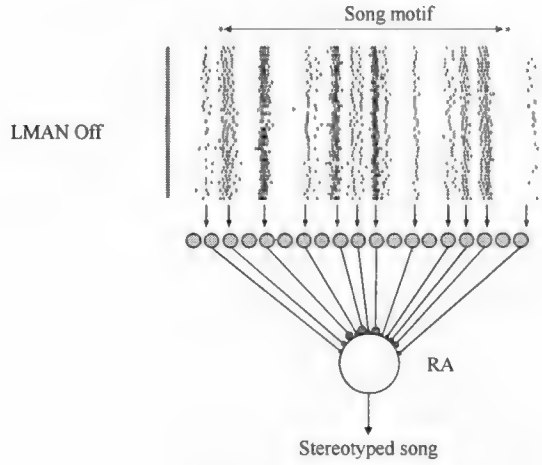
Shaping of the RA motor program in juvenile birds



Assumption: RA firing events are driven from HVC

Input from HVC is temporally precise

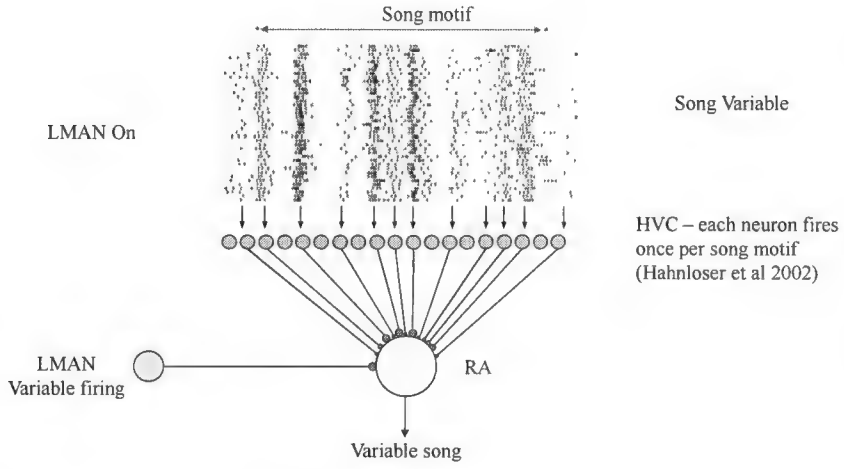
Shaping of the RA motor program in juvenile birds



Assumption: RA firing events are driven from HVC

Input from HVC is temporally precise

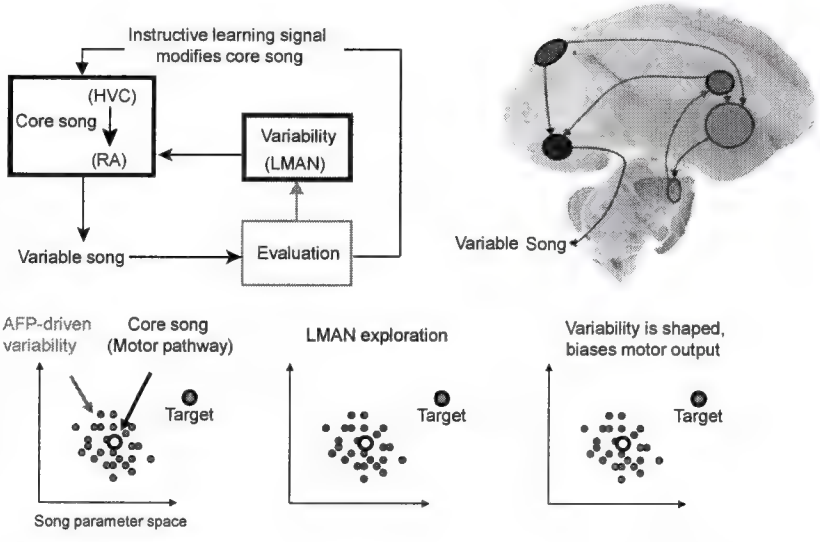
LMAN contributes variability to RA projection neurons



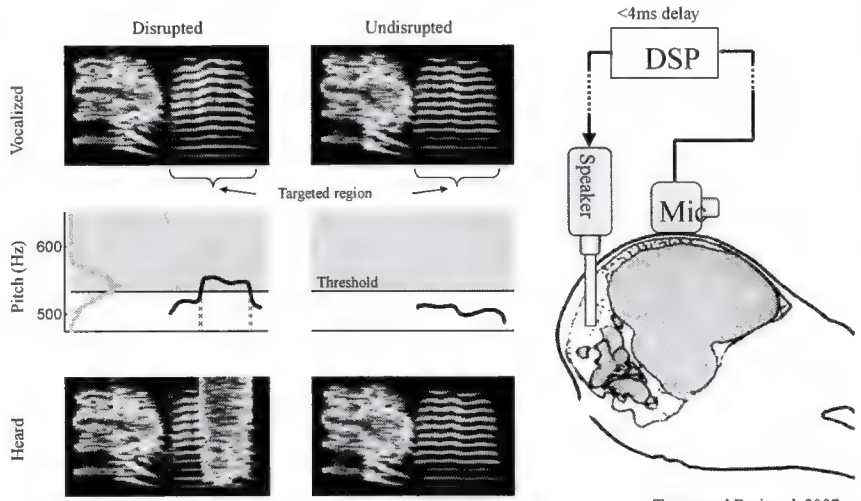
LMAN introduces variability in RA around the mean activity pattern driven by HVC

Is the variability produced by LMAN random or shaped?

Reinforcement learning algorithm

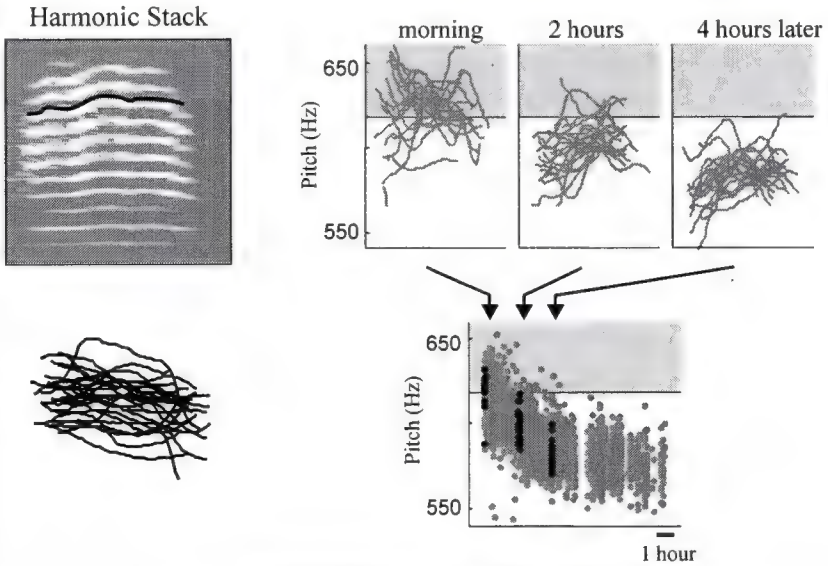


Rapid learning paradigm: Conditional Auditory Feedback



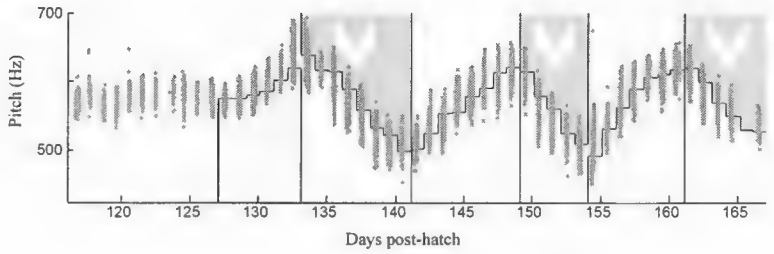
Turner and Brainard, 2007; Andalman and Fee, 2009

Conditional disruptive feedback results in rapid learning



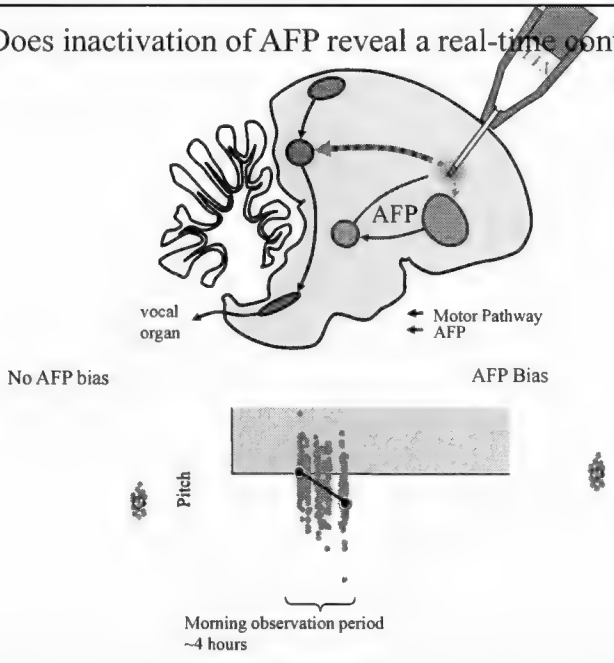
Turner and Brainard, 2007; Andalman and Fee, 2009

Conditional disruptive feedback results in rapid learning

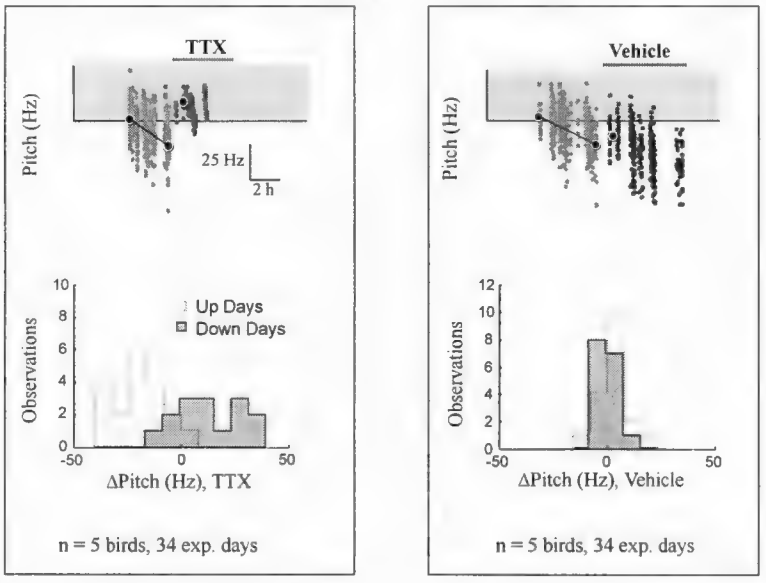


Andalman and Fee, 2009

Does inactivation of AFP reveal a real-time contribution?

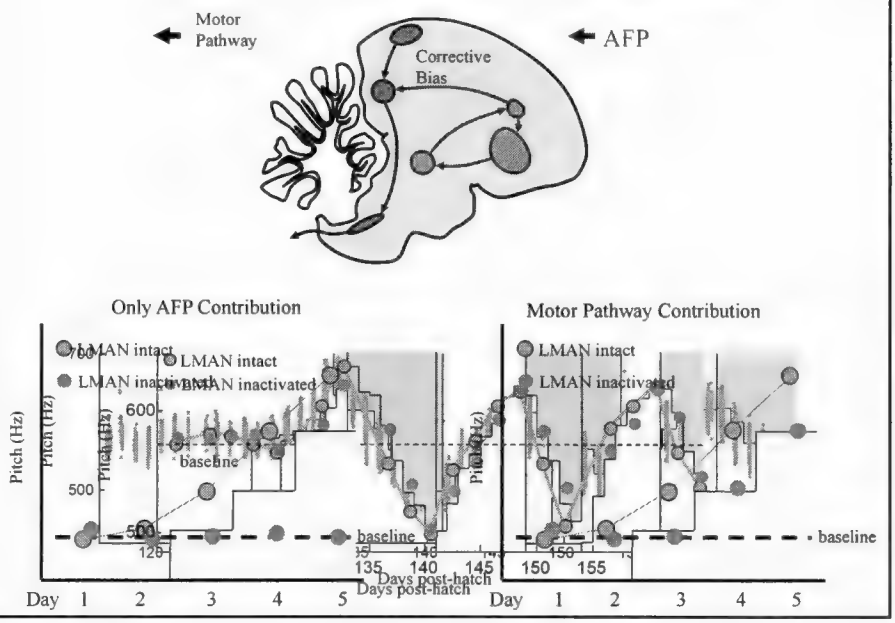


The AFP biases motor output to reduce auditory error



Andalman and Fee, 2009

What is the extent of the AFP contribution?



How does the motor program change as a function of learning?

Exploration → Habit

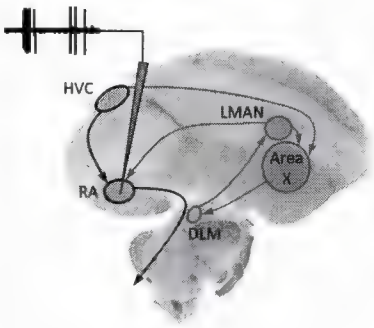
Variable motor performance

Large 'errors'

Stereotyped motor performance

Small 'errors'

Changes in the RA motor program during learning



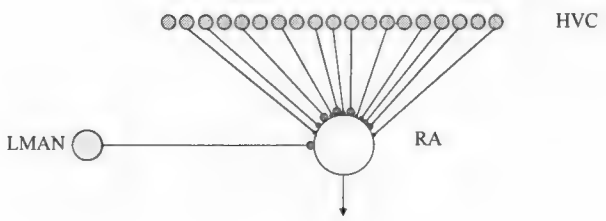
What changes during learning?

Spike trains become more bursty, sparse and stereotyped with learning

Changes in the song circuit during learning

HVC-RA synapses become fewer and stronger

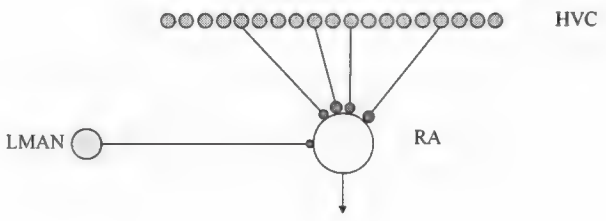
Herman & Arnold, 1991
Kittelberger & Mooney, 1999



Changes in the song circuit during learning

HVC-RA synapses become fewer and stronger

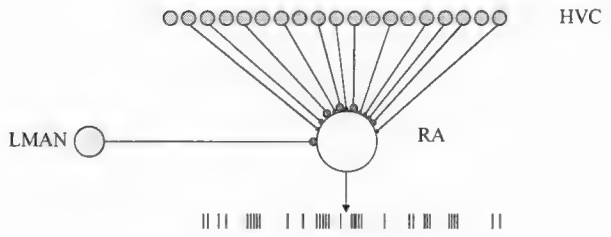
Herman & Arnold, 1991
Kittelberger & Mooney, 1999



Changes in the song circuit during learning

HVC-RA synapses become fewer and stronger

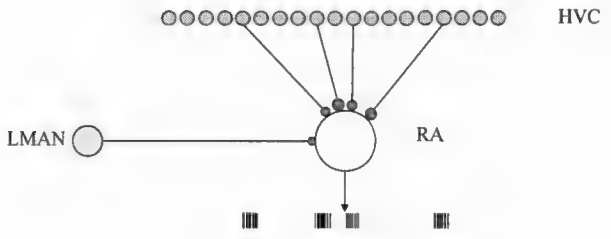
Herman & Arnold, 1991
Kittelberger & Mooney, 1999



Changes in the song circuit during learning

HVC-RA synapses become fewer and stronger

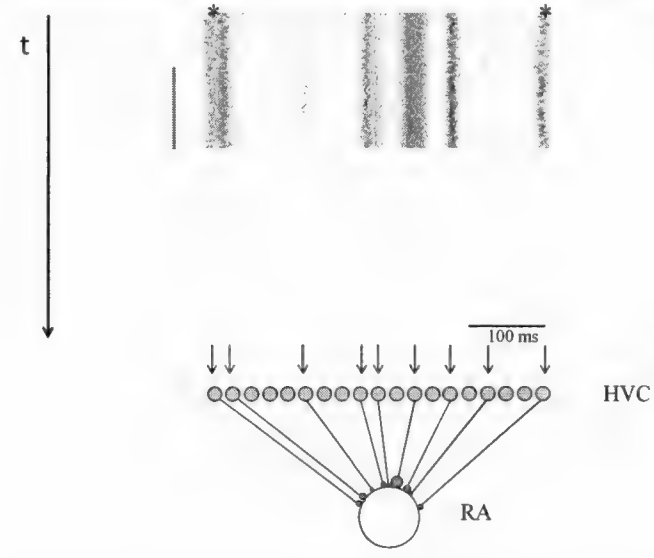
Herman & Arnold, 1991
Kittelberger & Mooney, 1999



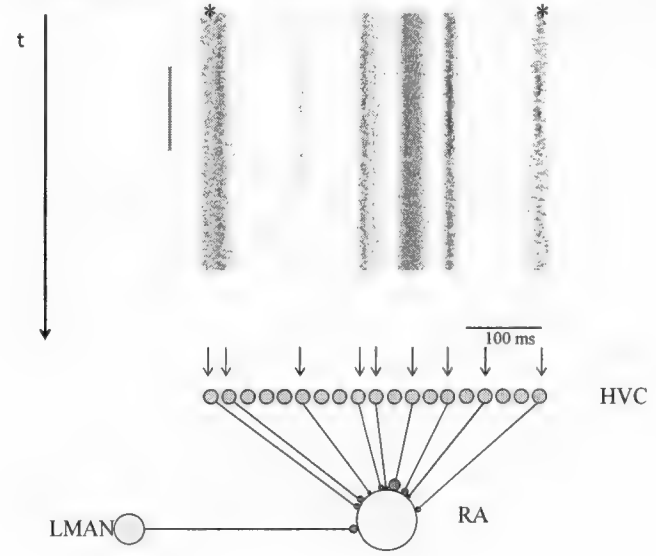
RA: Bursts become sparser and higher frequency

Firing patterns in RA reveal possible synaptic remodeling

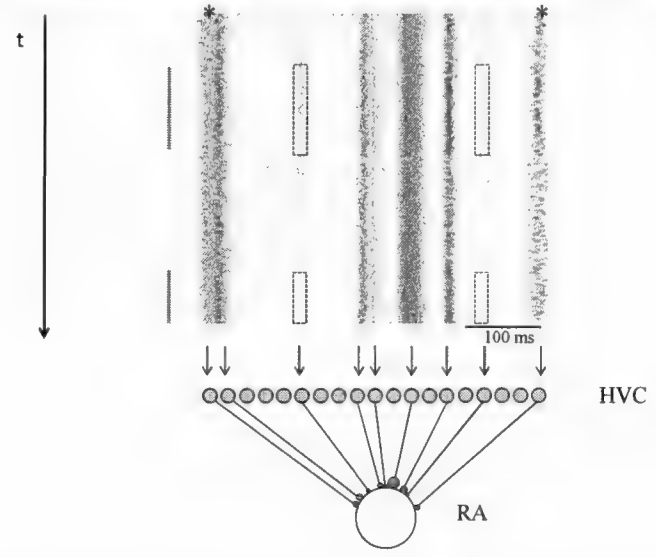
Firing patterns in RA reveal possible synaptic remodeling



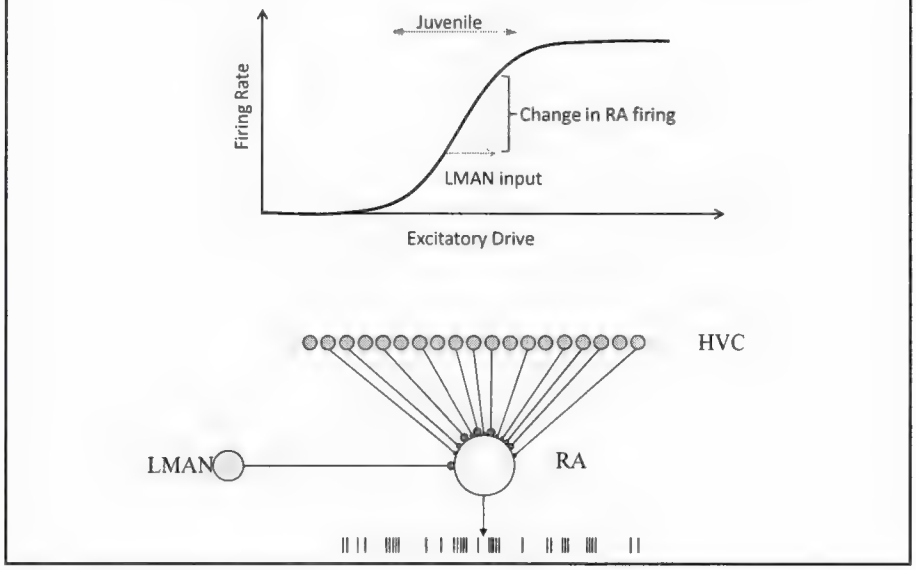
Firing patterns in RA reveal possible synaptic remodeling



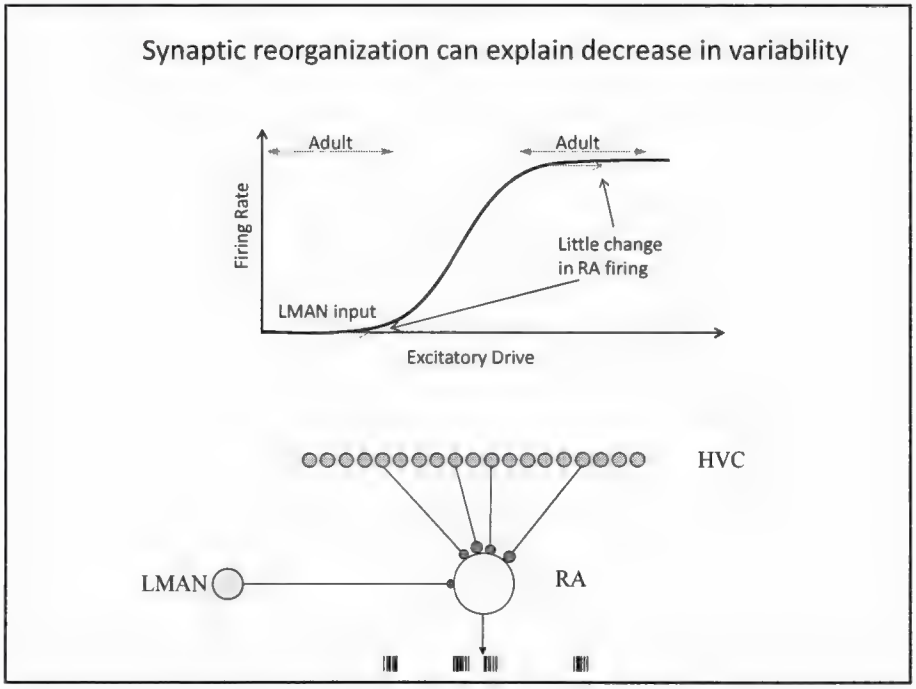
Firing patterns in RA reveal possible synaptic remodeling



Synaptic remodelling can explain decrease in variability

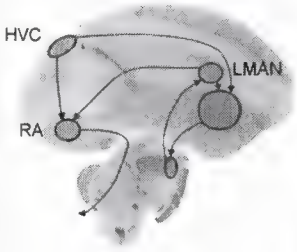


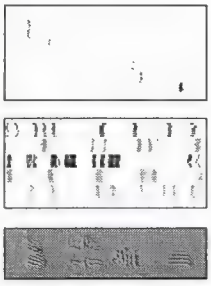
Synaptic reorganization can explain decrease in variability

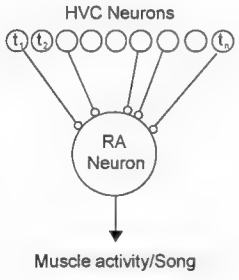


Summary

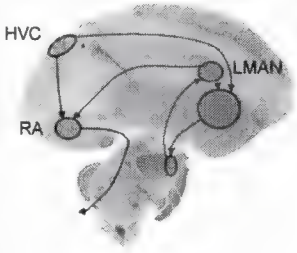
HVC represents the song as a time sequence. During learning connections between time-keeper neurons in HVC and muscle-related neurons in RA are formed such that the temporal sequence in HVC is transformed into a muscle activation patterns that produces the 'right' song.





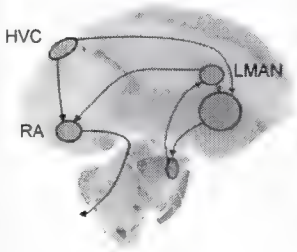


Summary



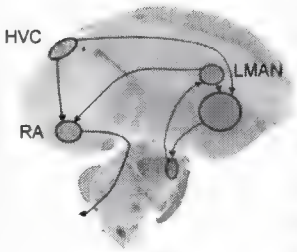
LMAN provides variable input to the motor pathway driving the vocal exploration needed for reinforcement learning.

Summary



The variability in LMAN can be 'informed', thus providing a corrective bias to the motor pathway. This AFP-produced bias slowly gets incorporated into the motor program.

Summary



Experience dependent synaptic reorganization –strengthening and pruning – at the HVC-RA synapse likely accompanies learning and serves to reduce the influence of LMAN, decreasing song variability.

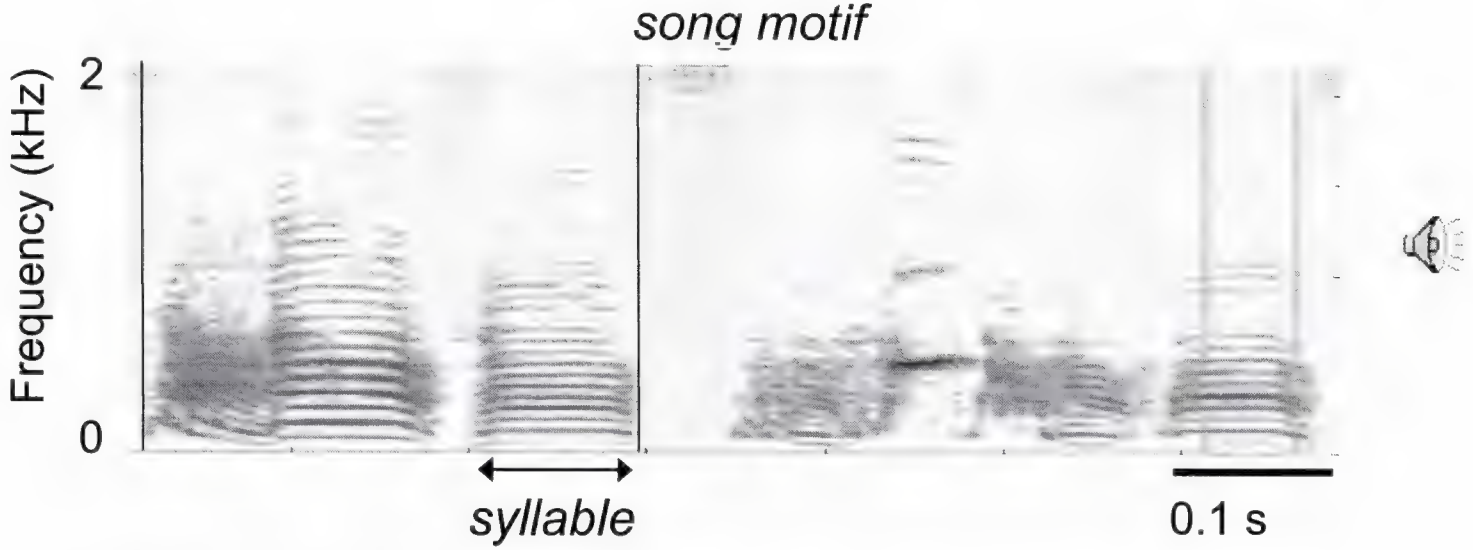
Birdsong learning as a microcosm of motor learning

MCN

MBL, 2009

Ila Fiete
Center for Learning and Memory, UT Austin.

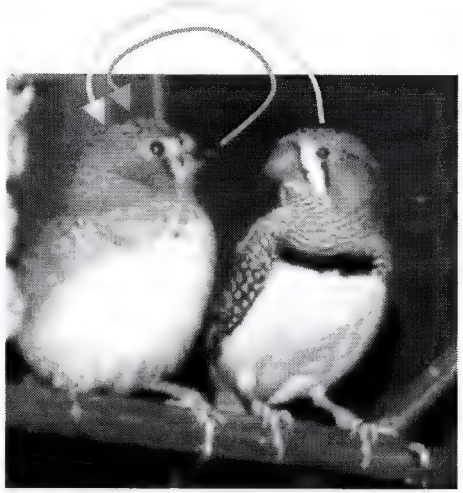
Birdsong

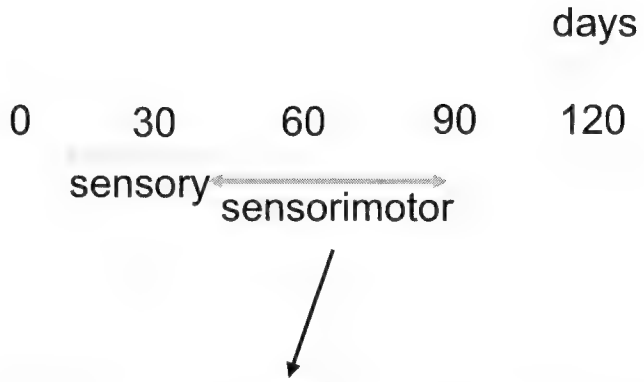


RHYMES WITH ORANGE by Hilary Price

THE FEEDER
YOU KNOW WHEN YOU GET A SONG STUCK IN YOUR HEAD AND YOU CAN'T STOP SINGING IT? ... I'VE HAD THE SAME SONG IN MY HEAD MY WHOLE STUPID LIFE.
HILARY B. PRICE 9-5rhymeswithorange.com

Behavior





Social feedback and tutor song not needed.

Auditory feedback of own song crucial.

learns to match tutor song through practice from internally stored template

Song learning: a “cartoon” model for motor learning

Sequence production.

Goal-directed learning.

Well-defined goal:

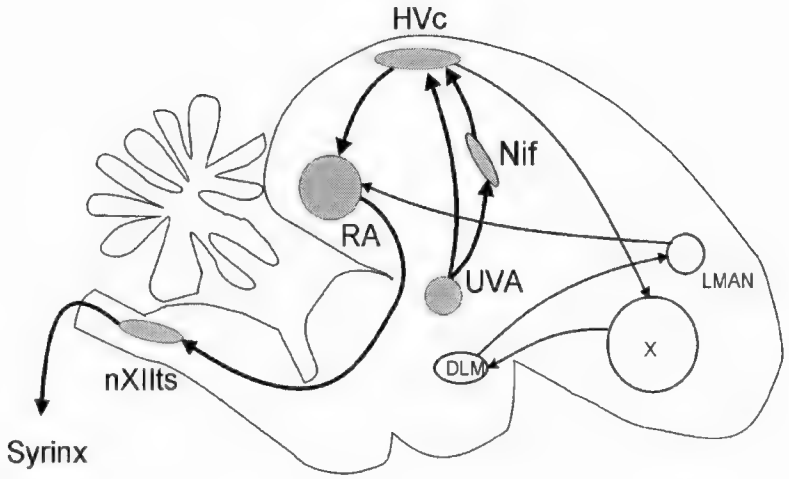
- single goal (tutor song) in some birds
- single set of syllables but flexible ordering in others
- no variable loads, etc.

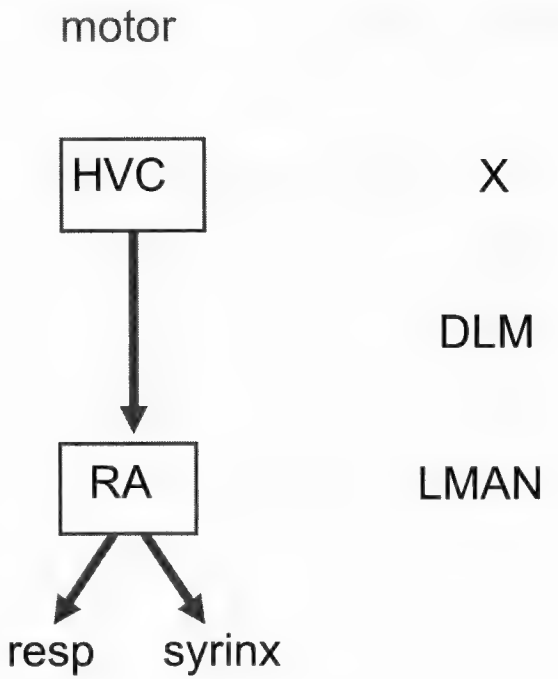
Neural pathways well-characterized.

Neural data from singing animals available.

microcosm for study of sequence production and goal-directed learning

Neural pathway for song



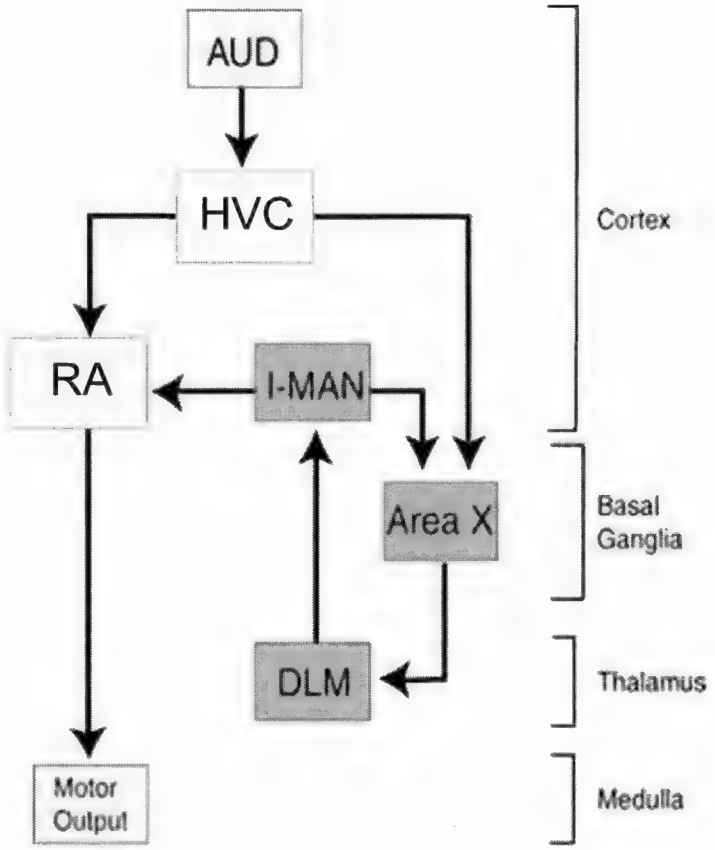


Lesion:

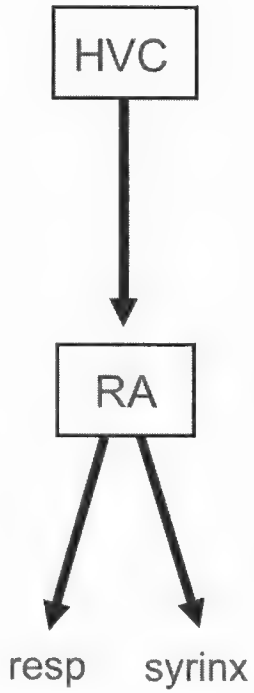
_____ no song

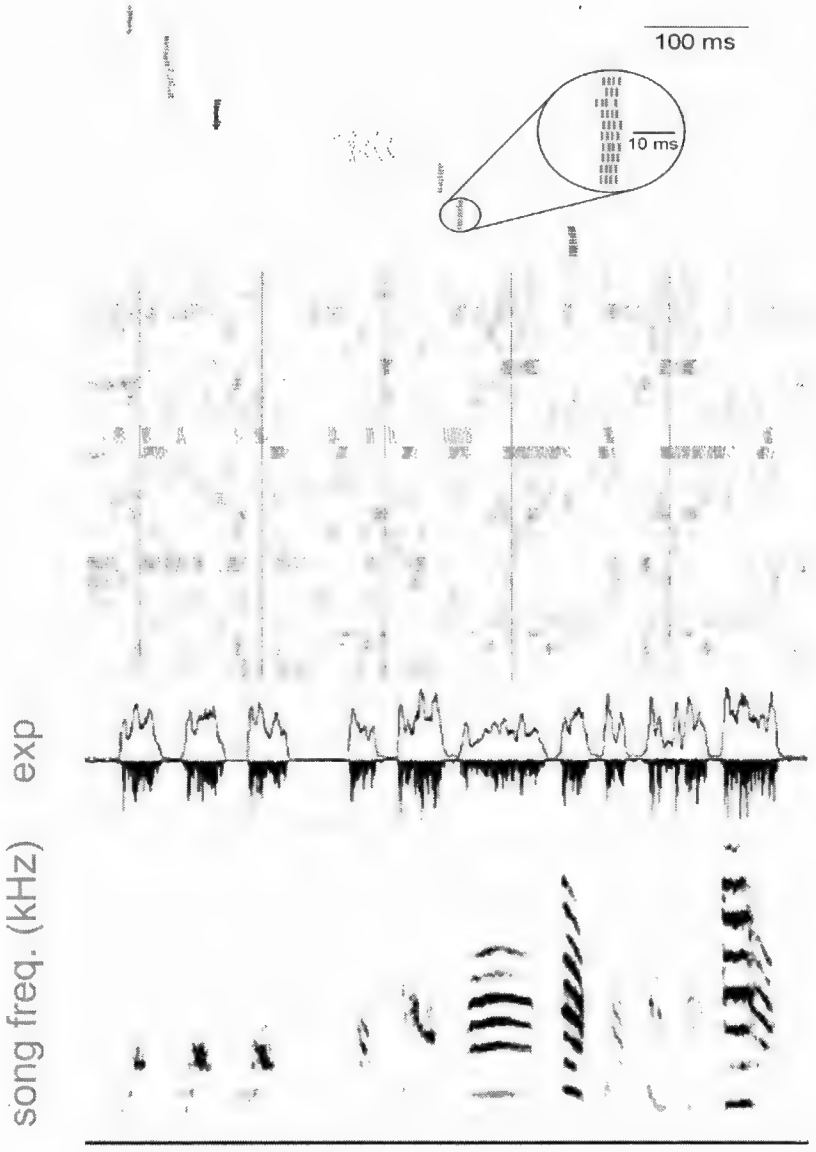
song unaffected

Song circuit



Activity in the premotor circuit



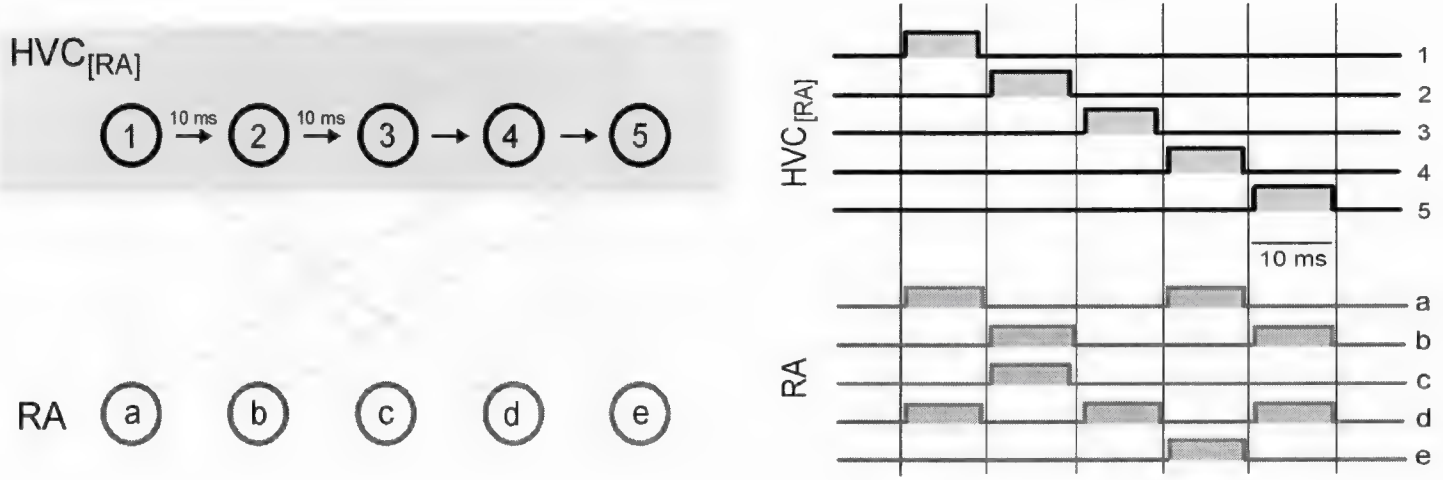


Hahnloser
et al., 2002

Leonardo
and Fee, 2005

Suthers
et al., 1999

Elements of song acquisition



I. HVC generates a generic, time-spooling sparse sequence.
II. Later, later stages learn the motor map specific to reproduce tutor song.

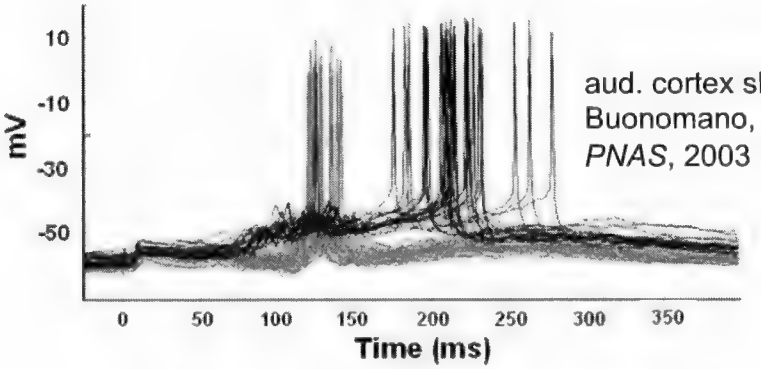
I. Formation of sequence-generating networks

"Whoa! *That* was a good one! Try it, Hobbs—just poke his brain right where my finger is."

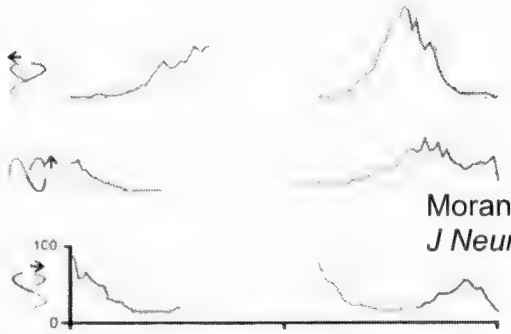
G. Larson, 1986

With W. Senn and R. Hahnloser

Neural activity sequences are ubiquitous.

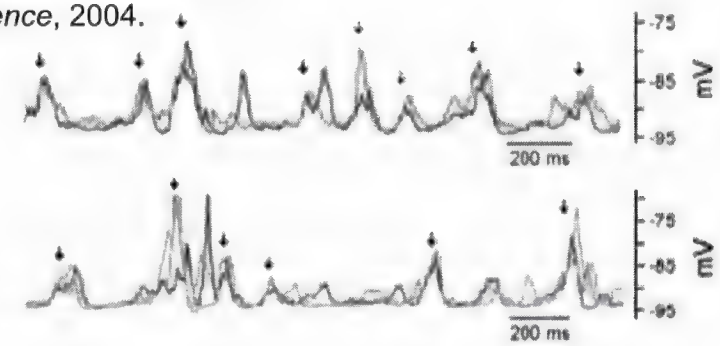


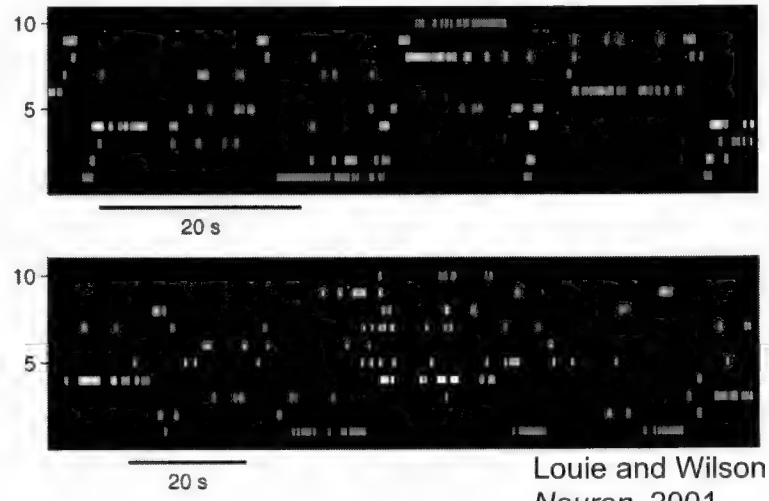
aud. cortex slice,
Buonomano,
PNAS, 2003



Moran & Schwartz,
J Neurophysiol, 1999

spont.,intracellular
Layer 5, Ikegaya et al.,
Science, 2004.





Louie and Wilson
Neuron, 2001

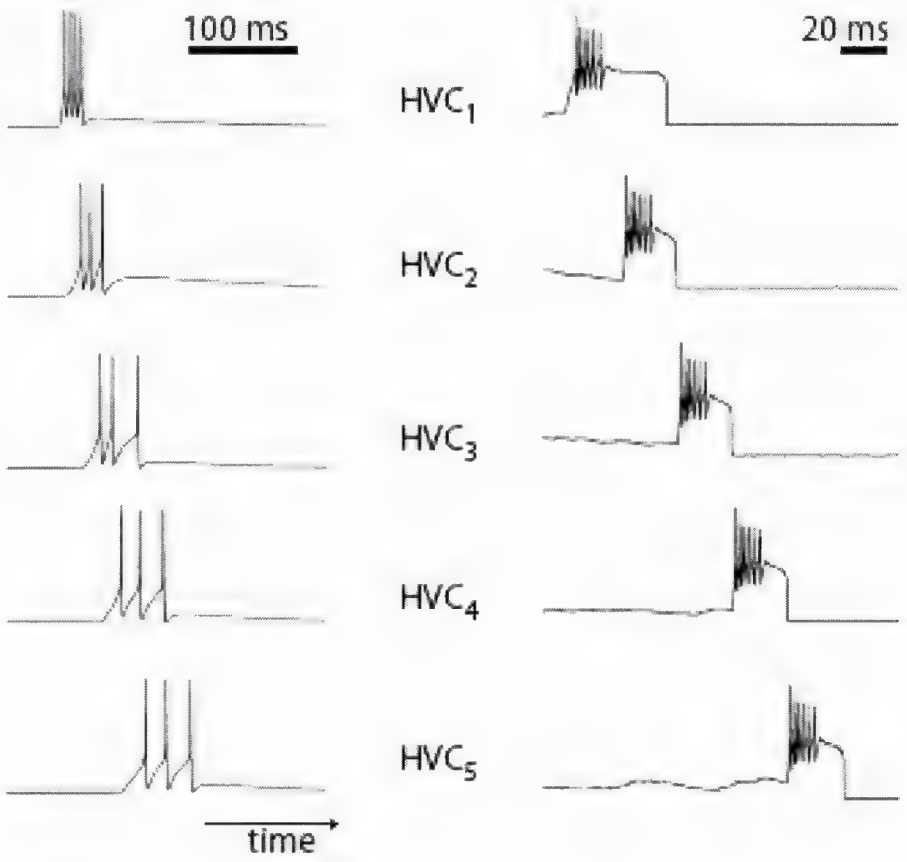
- Long compared to single-neuron activation.
- Stereotyped.

A few words about spatiotemporal pattern generation and synaptic chain networks (implications for HVC)

chalkboard

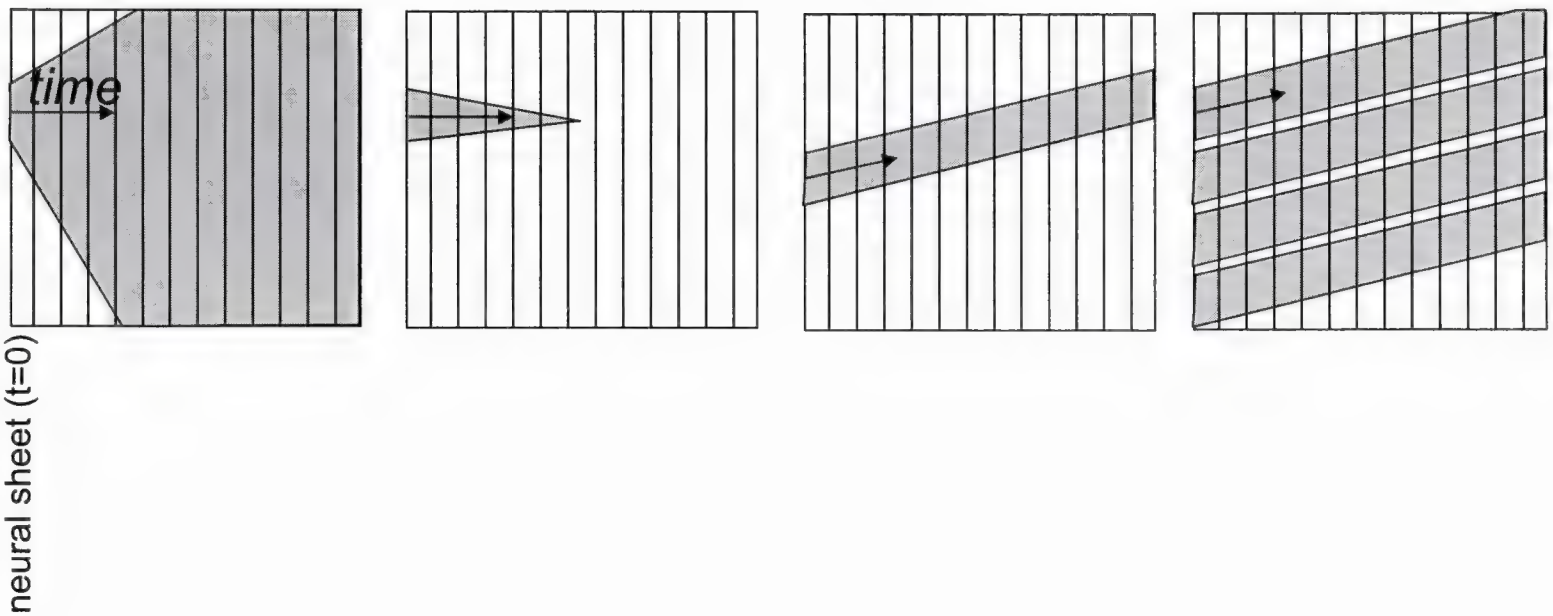
Lashley critique of associative chaining, 1951; Amari, 1972; Kleinfeld & Sompolinsky, 1988; Abeles, 1991; Li & Greenside, 2006, Jin et al., 2007; review: Fiete and Seung, New encyc. neurosci, 2008.

Sequence propagation in hardwired HVC models: role of bursts

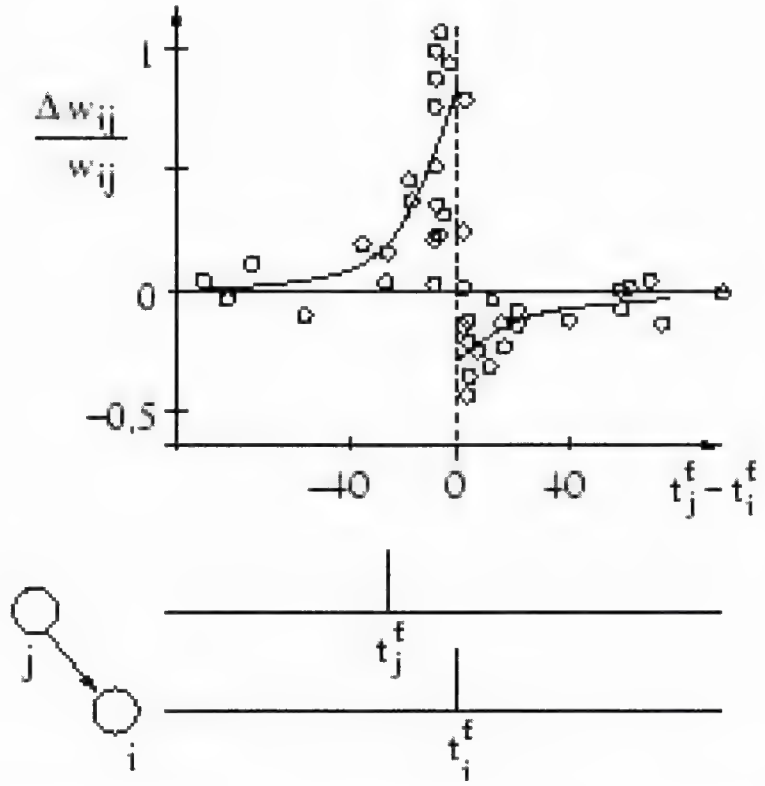


Li & Greenside, 2006, Jin et al., 2007.

Global space, time balancing to control activity propagation



Spike-time dependent plasticity (STDP)



Bi & Poo,
J Neurosci., 1998

repeated sequential input → predictive response

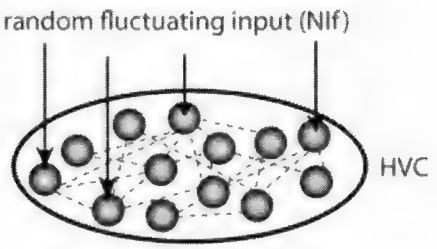
STDP and sequence learning

- Sequential training inputs
  - *Predictive* not *generative* responses. Rao & Sejnowski 2003
  - Require external clocked input for transitions. Drew & Abbott 2003
  - Synaptic delay, *tuned* global inhibition and adaptation, other fine homeostatic changes (e.g., threshold) necessary.
- Random inputs
  - Inequitable distribution of activity: synchronous bursts or very short sequences: ~5 steps ~10 ms, N=100-1000 neurons. Hertz & Prugel-Bennett, 1996; Levy et al., 2001

STDP → Activity bunching in space, time.

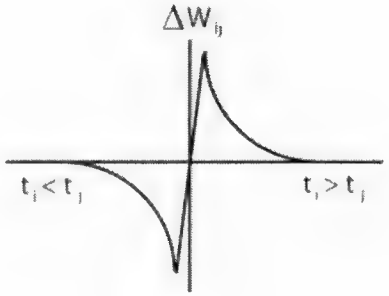
Hypothesis: local constraints on fan-in, fan-out could do it

0



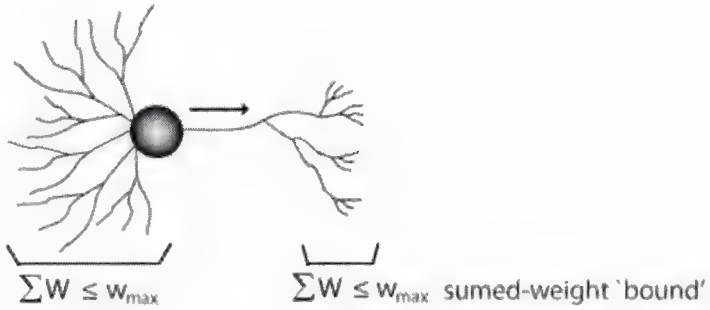
random network,
random inputs,
global inhibition

I



activity dependent
synaptic modification

II



limited synapse-building resources with heterosynaptic competition at each neuron

Summed weight limit rule

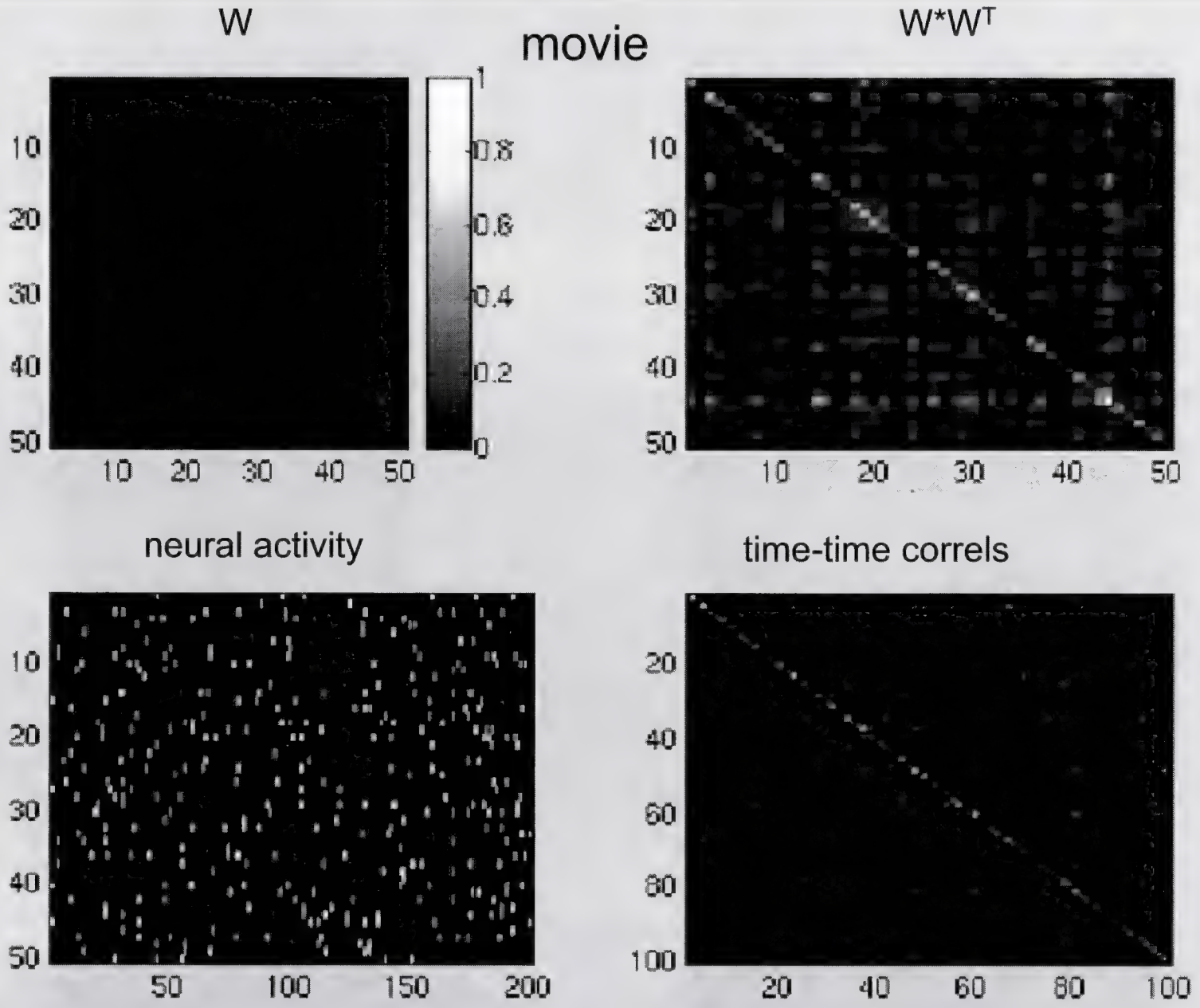
$$W_{ij} = W_{ij} + \eta \sum_{t,t'} x_i(t) K(t-t') x_j(t') - \varepsilon\eta \sum_k (W_{ik}\theta_{i*} + W_{kj}\theta_{*j})$$

$$\theta_{i*} = \Theta\left(\sum_k W_{ik} - W_{\max} \right) \quad \longleftarrow \text{summed wts to i}$$

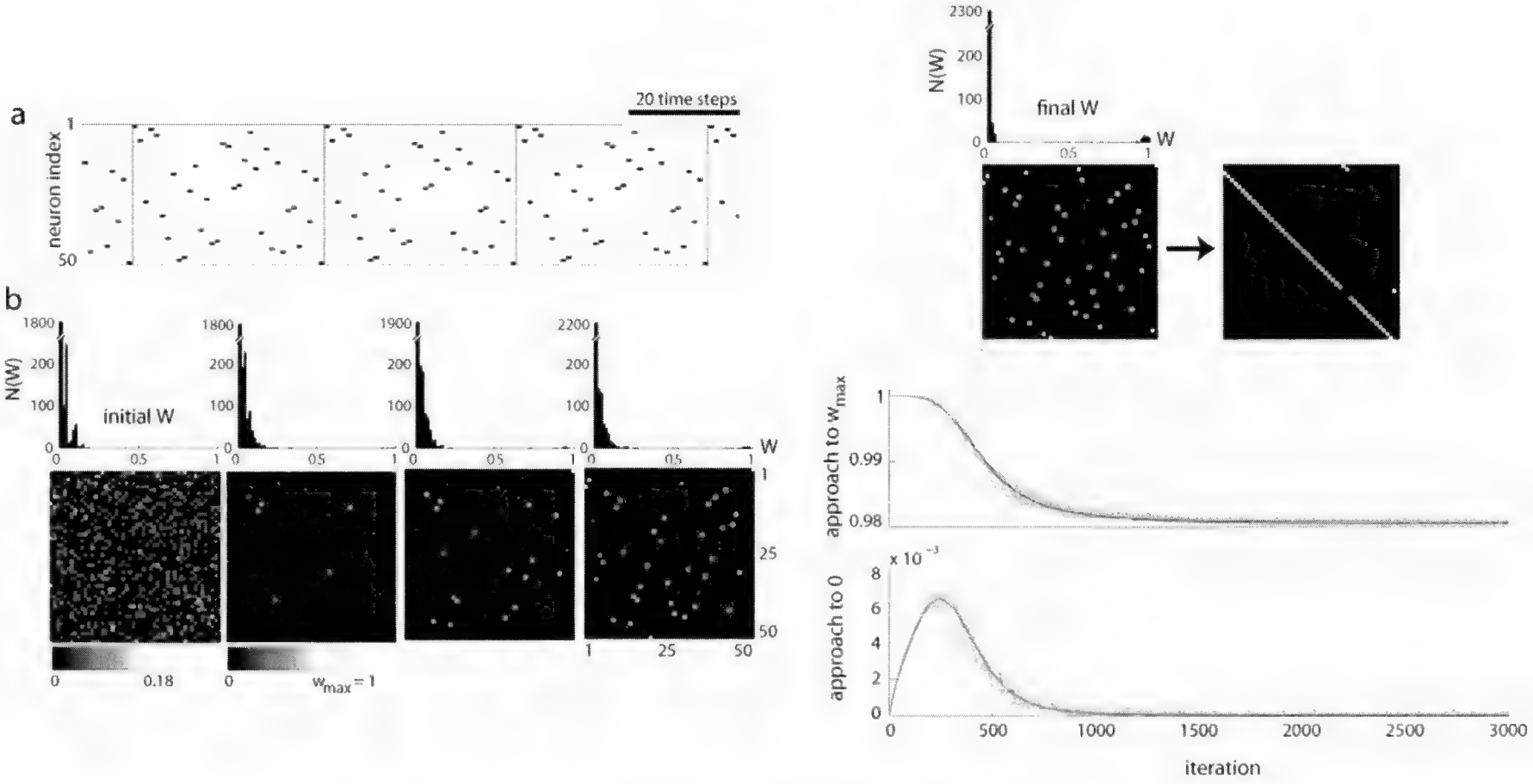
$$\theta_{*j} = \Theta\left(\sum_k W_{kj} - W_{\max} \right) \quad \longleftarrow \text{summed wts from j}$$

$$W_{ij} \in [0, w_{\max}] \quad \longleftarrow \text{hard individual weight bound}$$

$$m w_{\max} = W_{\max}$$

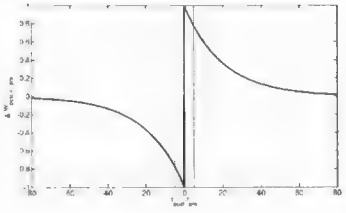

W
movie
W*W^T
10
20
30
40
50
1
0.8
0.6
0.4
0.2
0
neural activity
time-time correls
50
100
150
200
60
80

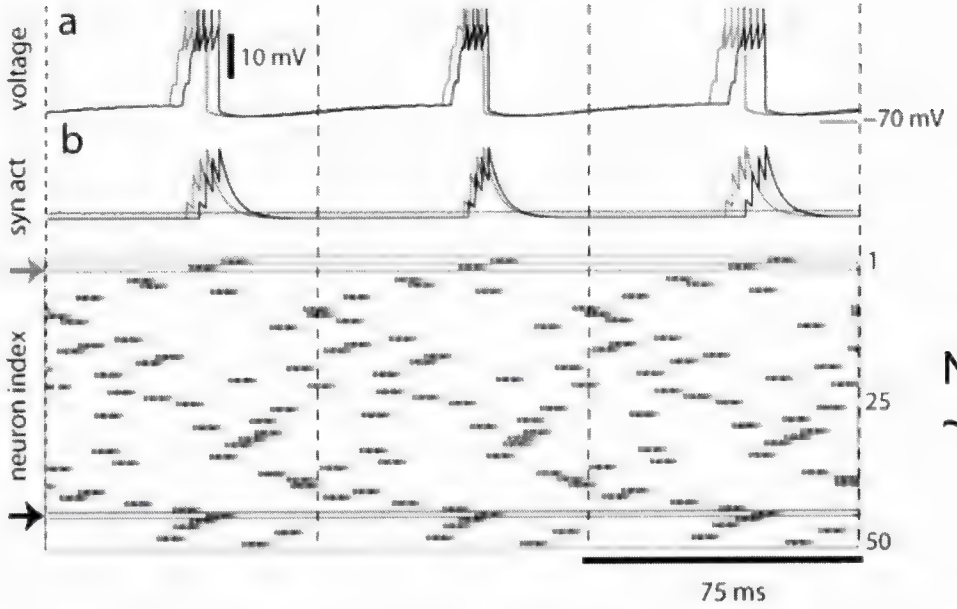
Summed weight limit



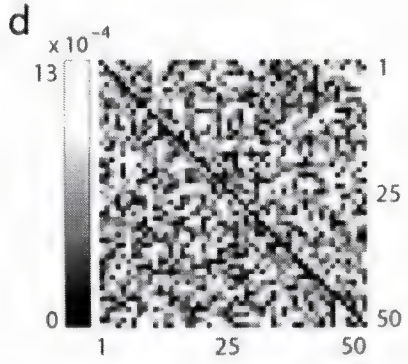
random weight matrix → permutation matrix!

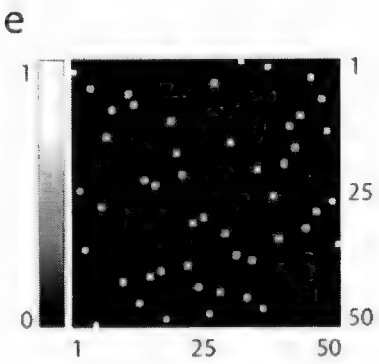
Conductance-based neuron results



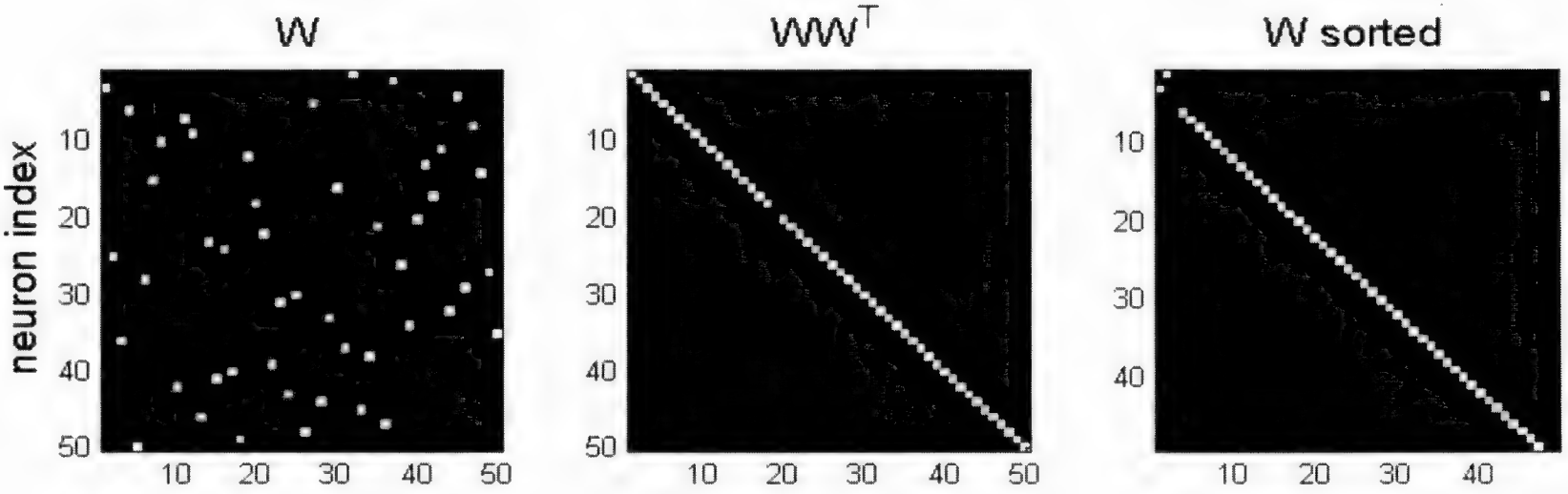


N=50, T~80 ms:
~10 'steps'



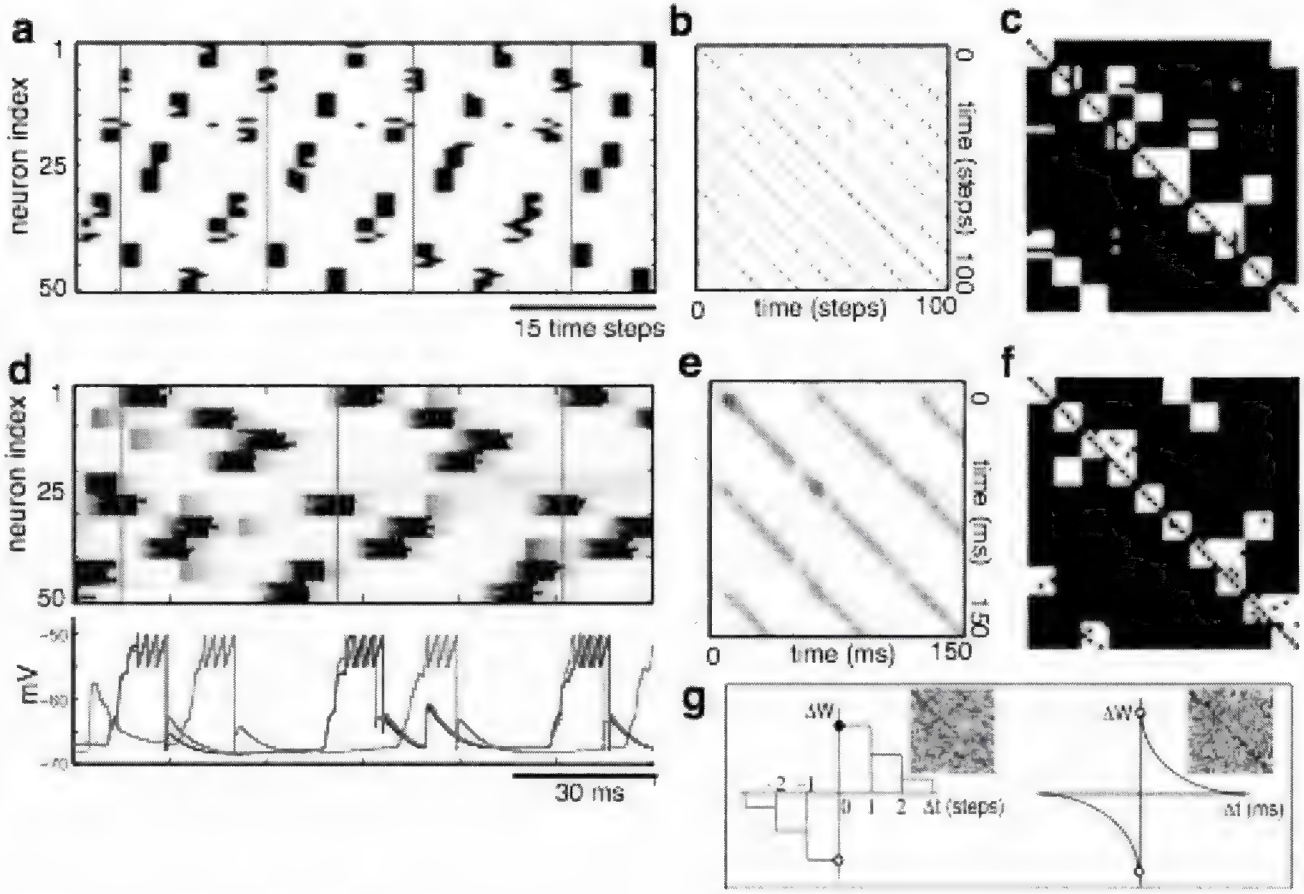


I&B spiking network



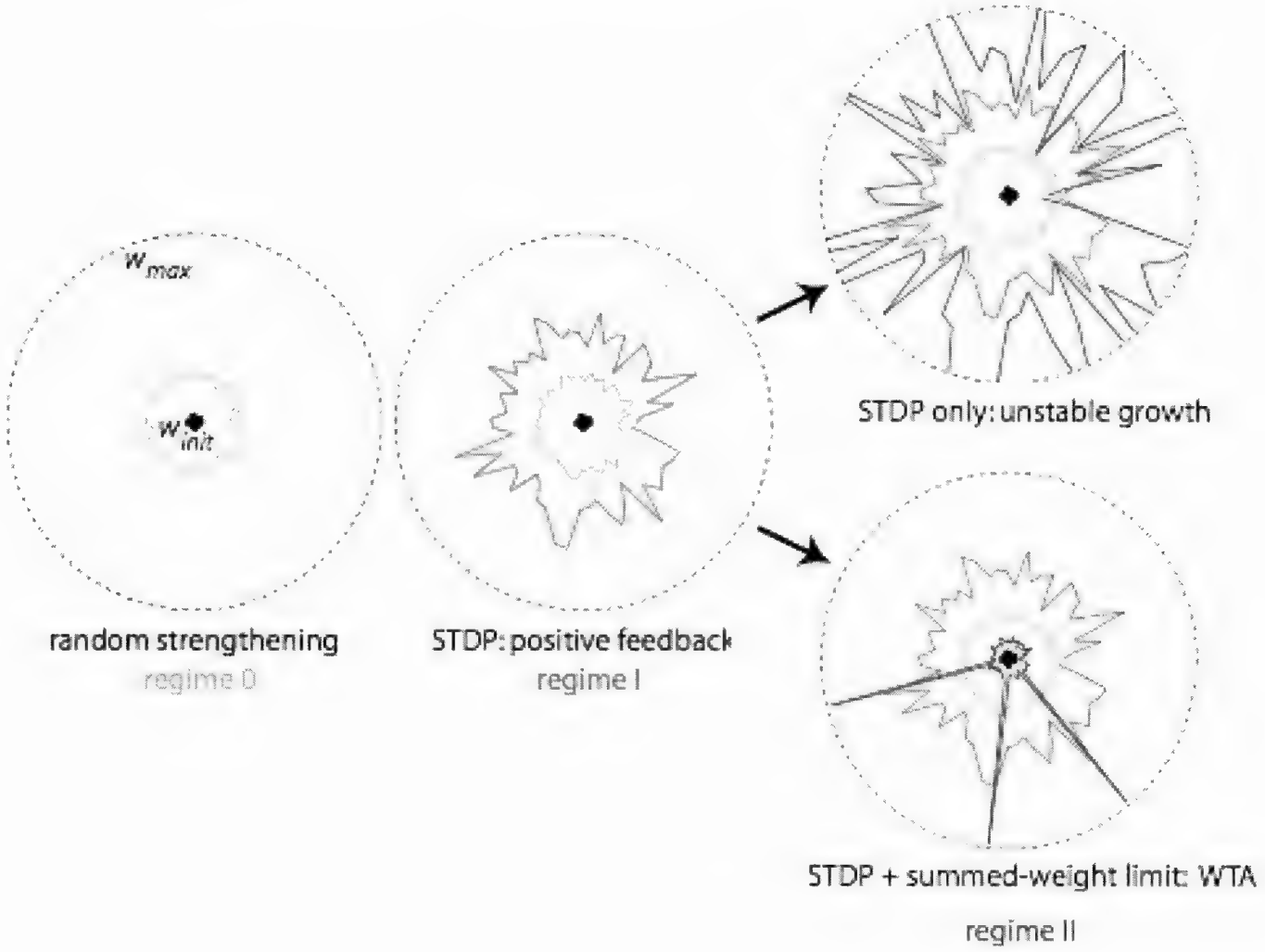
Weight matrix → permutation matrix

Wide chains (m>1)



disjoint inputs groups (k neurons each), stochastic synapses, neural response,

Heuristic mechanism: 3 regimes



Reduced model suggests uniformly sampled permutation matrices: scale free.

Selective chain growth, as a function of chain length?

Any other dynamics or time-scales?

Toy model: chain-length distribution

ways to get a length L chain:

$$\frac{(N \cdot N-1 \cdot N-2 \cdots (N-L+1) \cdot 1)}{L} \cdot (N-L)!$$

select L elements for chain

remaining elements

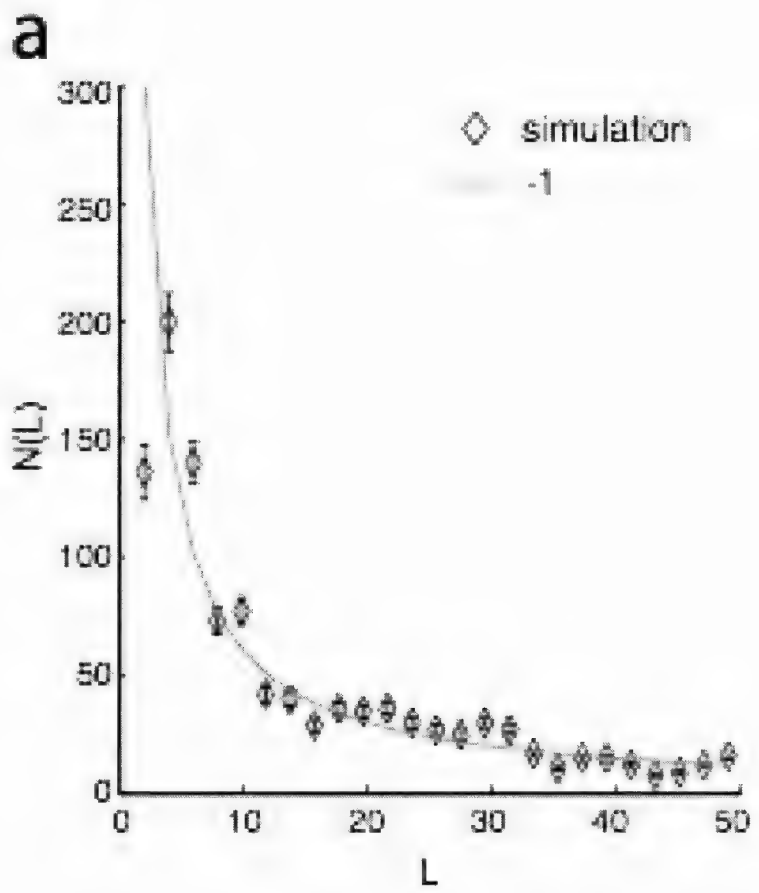
total # of arrangements:

$$N!$$

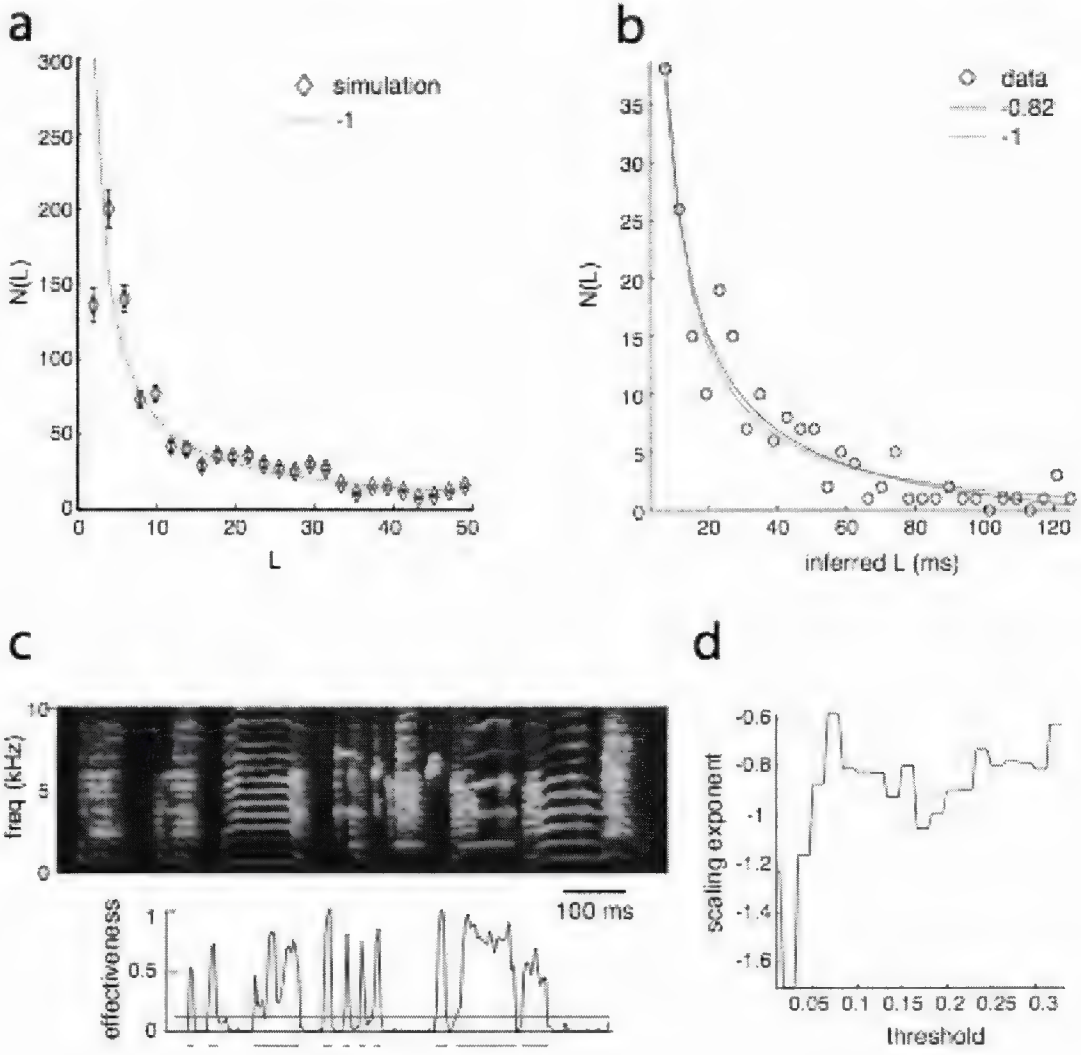
probability of a length L chain:

$$p(L) = \left(\frac{N!}{L}\right) \cdot \left(\frac{1}{N!}\right) = \frac{1}{L}$$

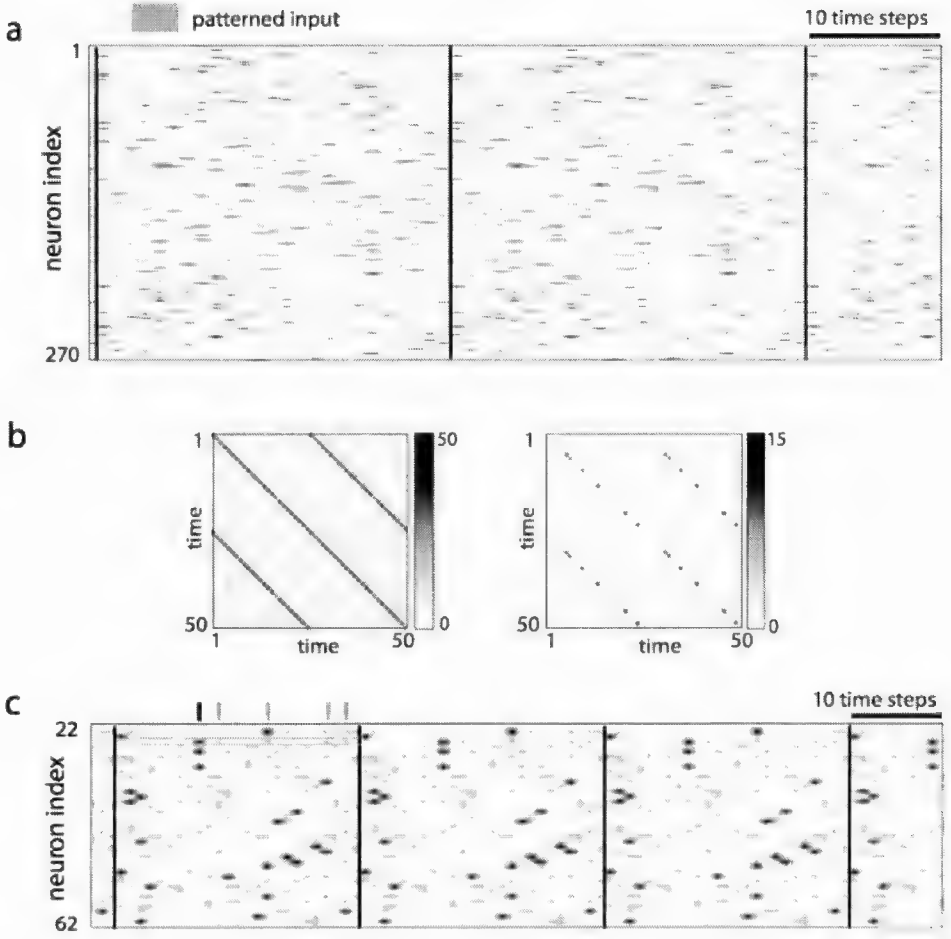
Heuristic model and simulated length distributions



Actual and predicted length distributions

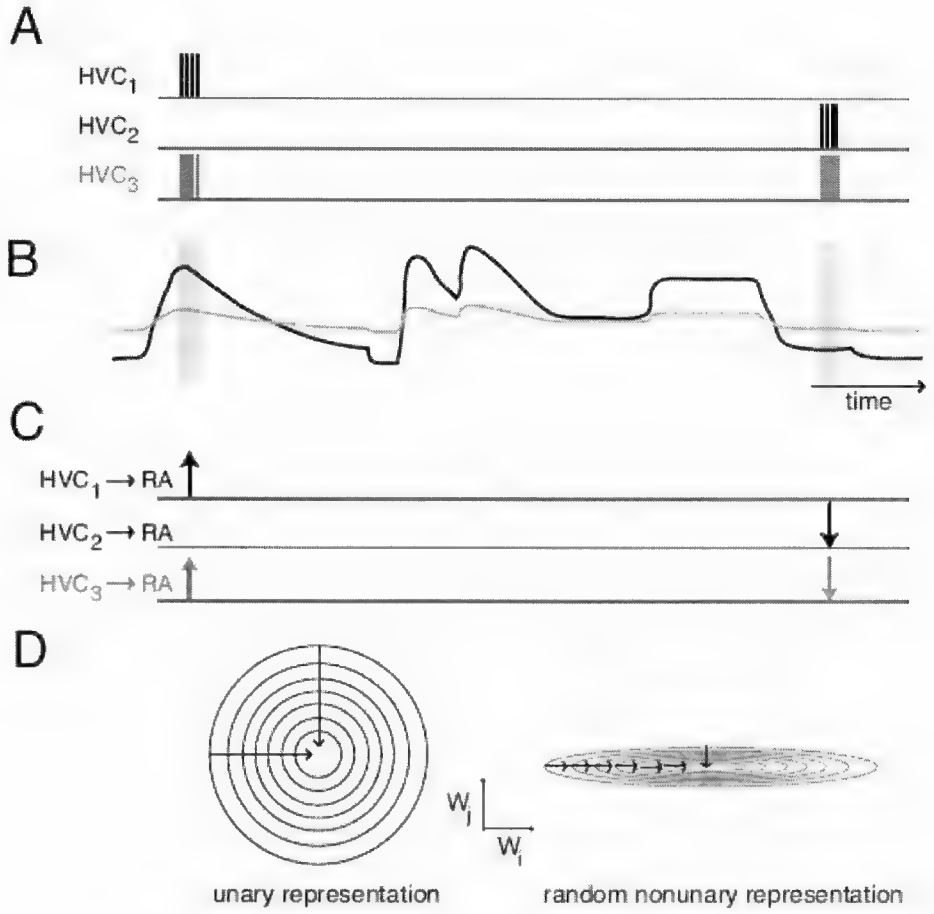


Sequential inputs



Dense input sequence of random patterns sparsified.
Extremely fast learning (~10-50 iterations).

Why sparse sequence?



Fast learning of forward motor map: minimize interference of synaptic changes for learning different parts of song.

Summary

- Predict existence of heterosynaptic competition for synapse building: heterosynaptic LTD following STDP-induced LTP in any synapse. Bonhoffer, Royer and Pare.

- Chain formation can occur even with random inputs, but time-course slower than with sequential input.

- Chain length distribution is scale-free power law (with -1 exponent) when network is formed with non-sequential inputs.

- Underlying synaptic chain architecture: block permutation matrices.

- Generic and robust synaptic chaining mechanism.

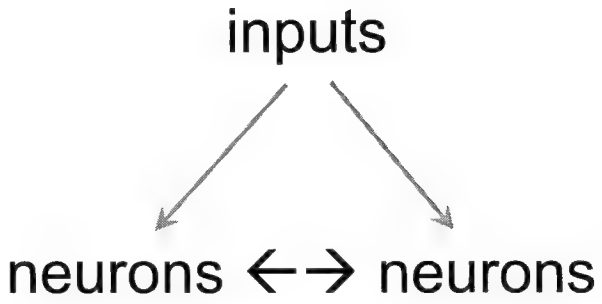
II. Goal-directed learning

Define goal: Learn the reward schedule of the world
(e.g. TD learning, dopamine).

Achieve goal: Learn which actions lead closer to goal, and learn to perform these actions (e.g. REINFORCE, ?).

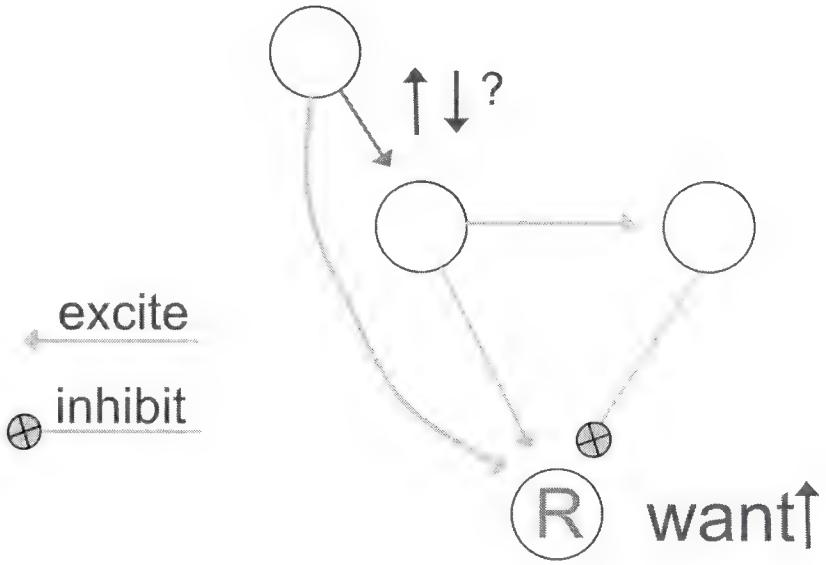
	define goal	achieve goal
reward optimization	learn reward schedule, (output → reward mapping) predict reward	learn how to produce outputs which maximize reward
machine learning	TD learning	REINFORCE
brain	dopamine system, hebbian learning	actor network, ?

Hebb

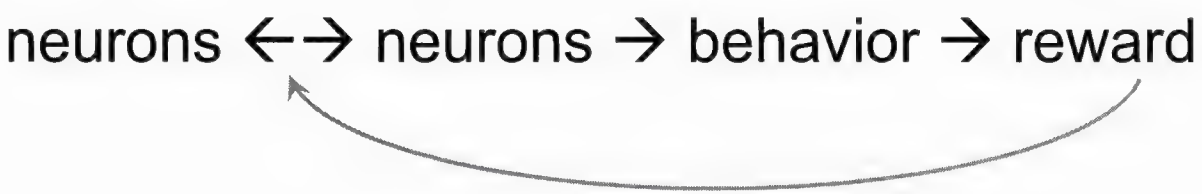


coincident firing → strengthen connection

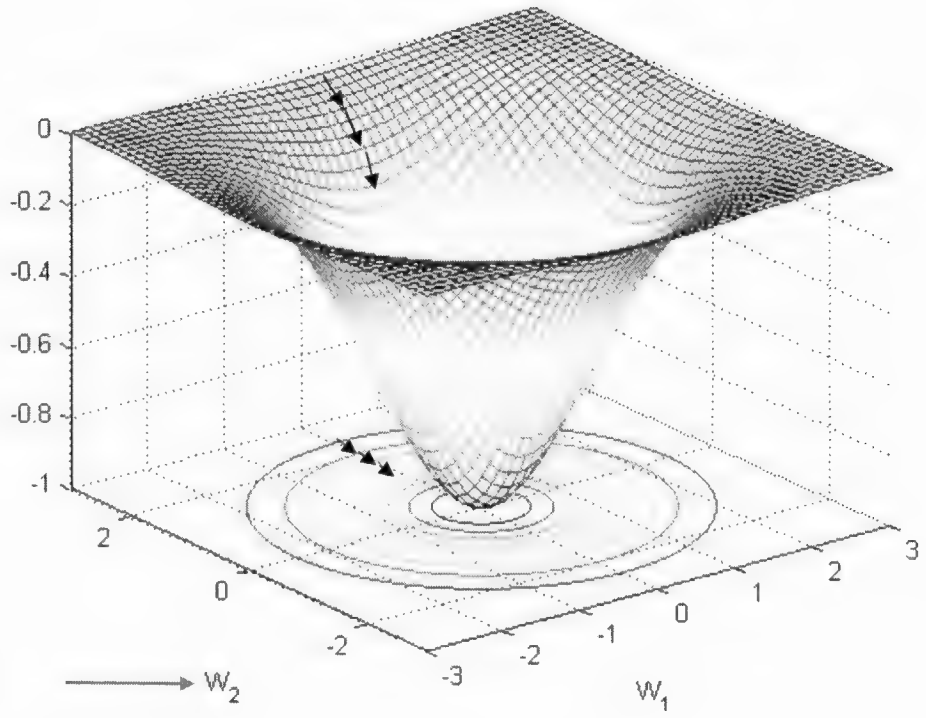
Activity-reward correlations (Hebbian learning) insufficient



Neural basis of reinforcement learning?



Gradient learning



synaptic weight parameters

$R_W[\Omega]$;

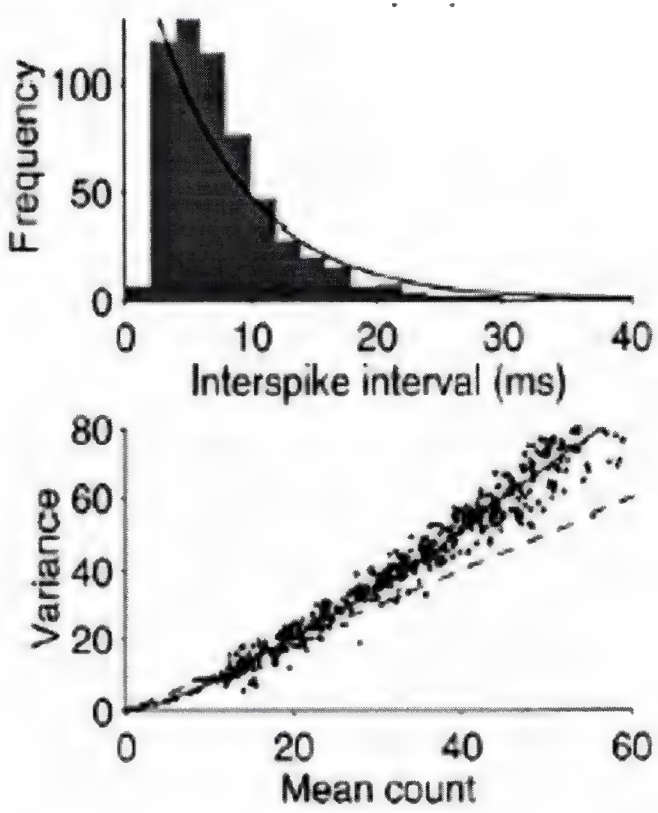
$$\boxed{\Delta W_\alpha = \frac{\partial R}{\partial W_\alpha}} \quad \Rightarrow \quad \Delta R > 0$$

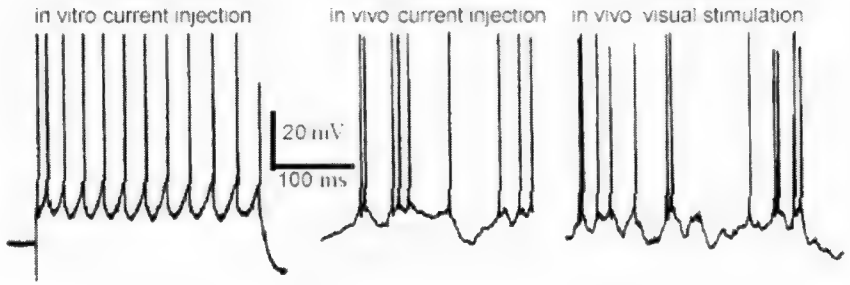
	random variation →	consolidation
evolution	mutation recombination	replication
bacterial chemotaxis	tumbling	tumbling suppression
learning	noise injection stochastic synapses*	synaptic strengthening

Barto and Anandan 1985/Williams 1992; Xie and Seung 2004, Fiete and Seung 2006.

*Minsky 1954; Barto 1983; Williams 1992; Seung 2003.

In vivo cortical firing is noisy





Shadlen and Newsome, 1998

Learning through individual neuron trial-and-error, and concerns about learning time

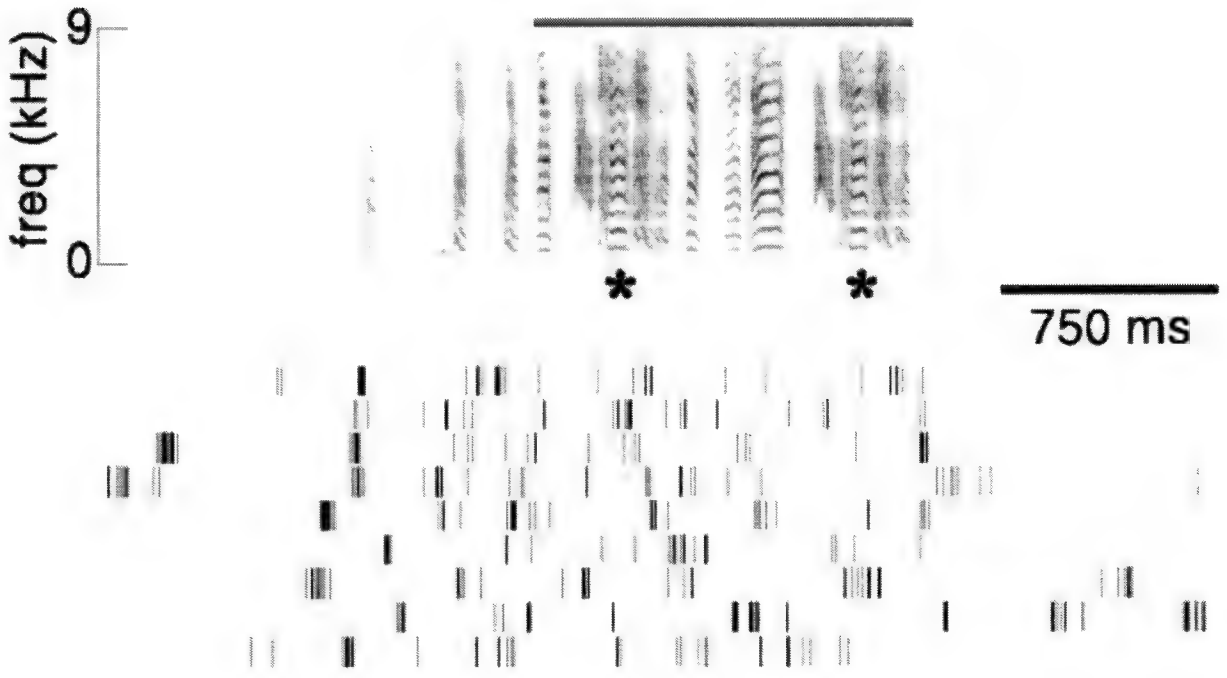
chalkboard

Barto and Anandan 1985; Williams 1992; Xie and Seung 2004; Minsky 1954; Barto 1983; Williams 1992; Seung 2003; Fiete and Seung 2007

Learning through individual neuron trial-and-error, in the songbird

With M. S. Fee and H.S. Seung

LMAN activity is variable



Hessler & Doupe, 2001

LMAN lesions → failure to learn song

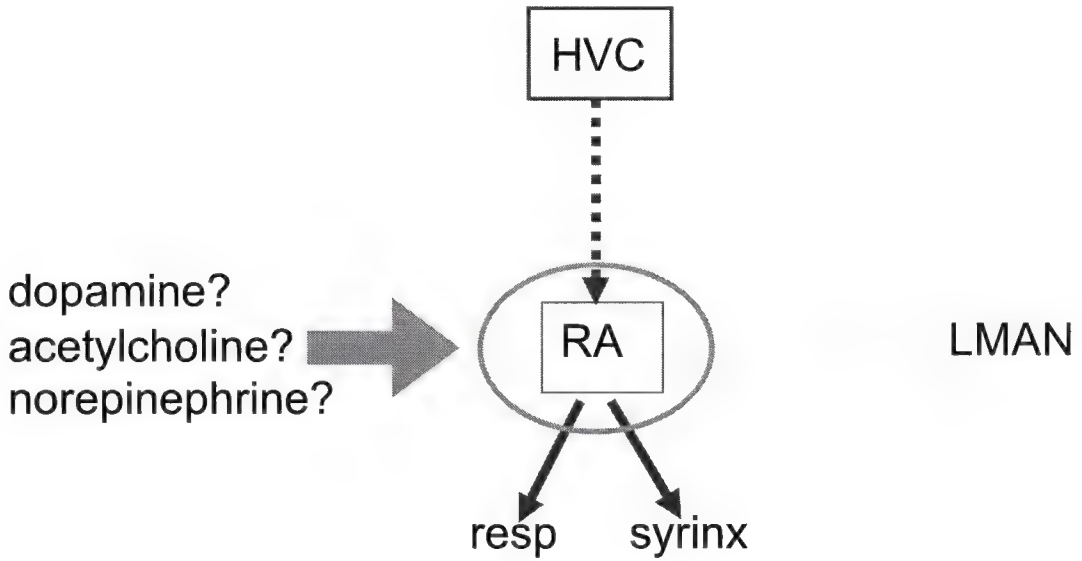
Scharff and Nottebohm, 1991

LMAN lesions → failure to learn song
low song variability
premature crystallization

Scharff and Nottebohm, 1991

LMAN hypothesis

LMAN drives conductance fluctuations in RA



Doya & Sejnowski 1995: LMAN perturbs HVC-RA weights through slow neuromodulation

Fiete and Seung 2003: LMAN rapidly perturbs RA neurons through regular synaptic transmission

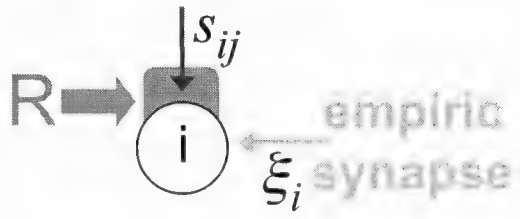
learning rule

$$\Delta W_{ij} = \eta R \cdot e_{ij} \qquad e_{ij} = \int dt \xi_i(t) s_{ij}(t)$$

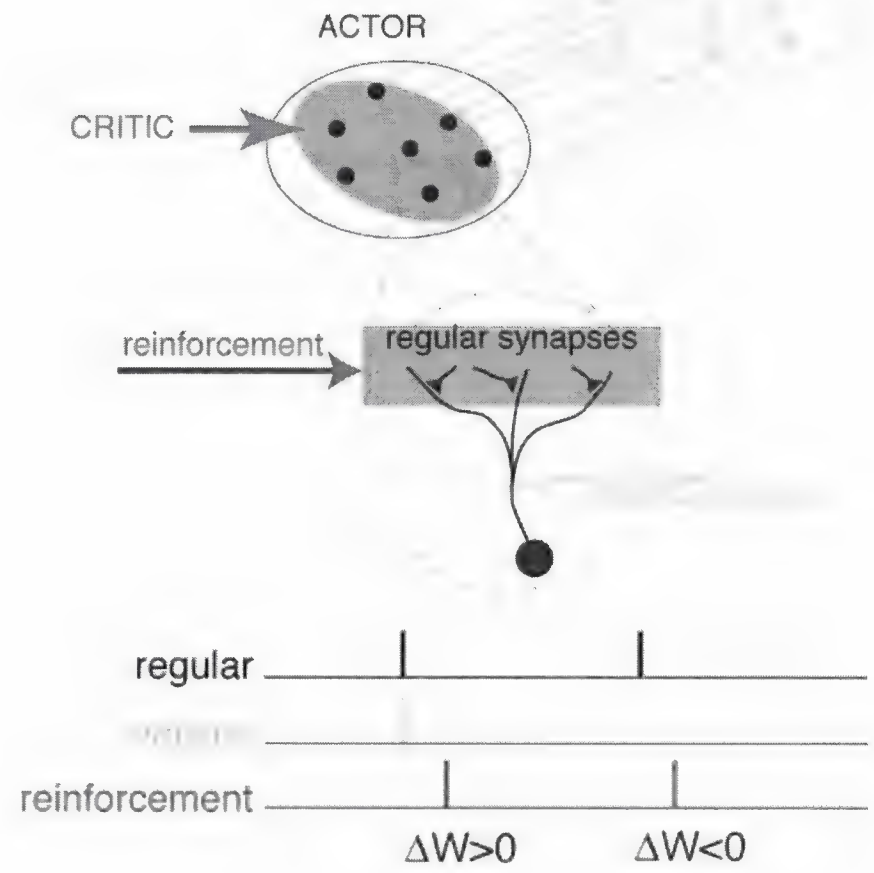
local synaptic eligibility trace

$$\left\langle \Delta W_{ij} \right\rangle \propto \frac{\partial \left\langle R \right\rangle}{\partial W_{ij}}$$

regular synapse



Synaptic learning rule



But LMAN experimenter or supervisor?

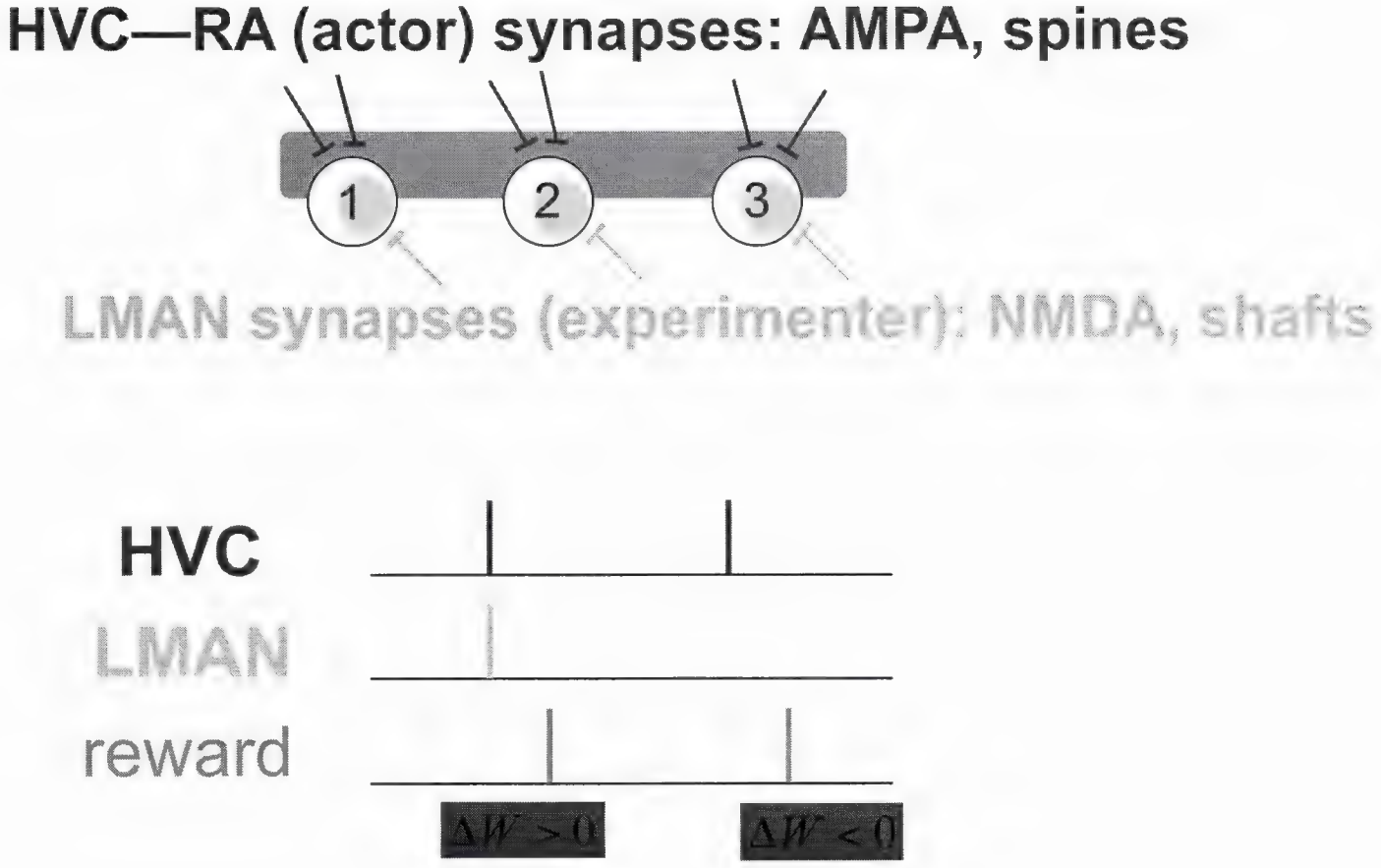
Instruction to postsynaptic neuron

- Increase in activity compared to normal (empiric input) followed by reward: “increase activity”
- Lower activity (quiescence of empiric input) followed by reward: “decrease activity”

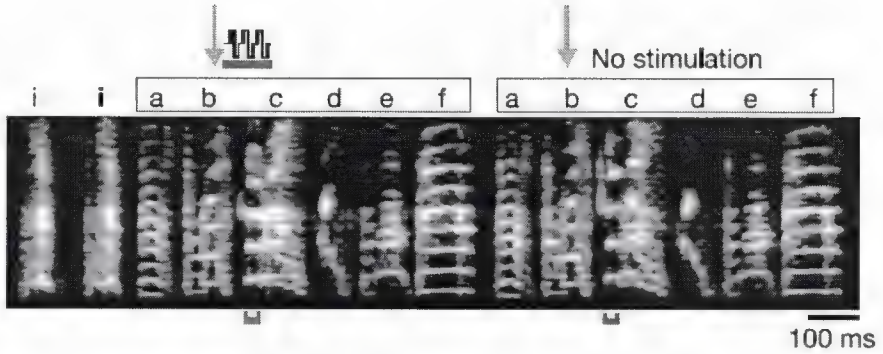
Stochastic gradient ascent

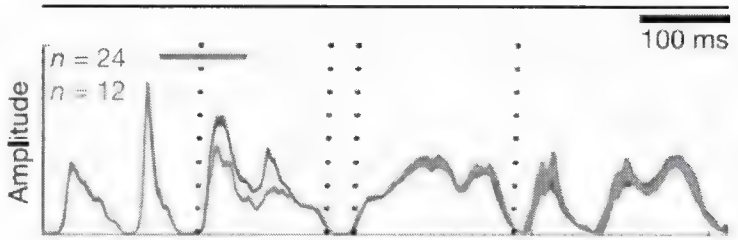
- Moves along gradient of reward
- Depends on local activity (+ global reward)
- Dynamically perturb N neurons to train N^2 static weights
- Model-free

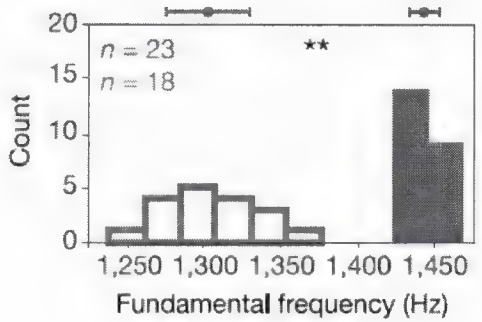
Cellular mechanism?



LMAN locally perturbs song

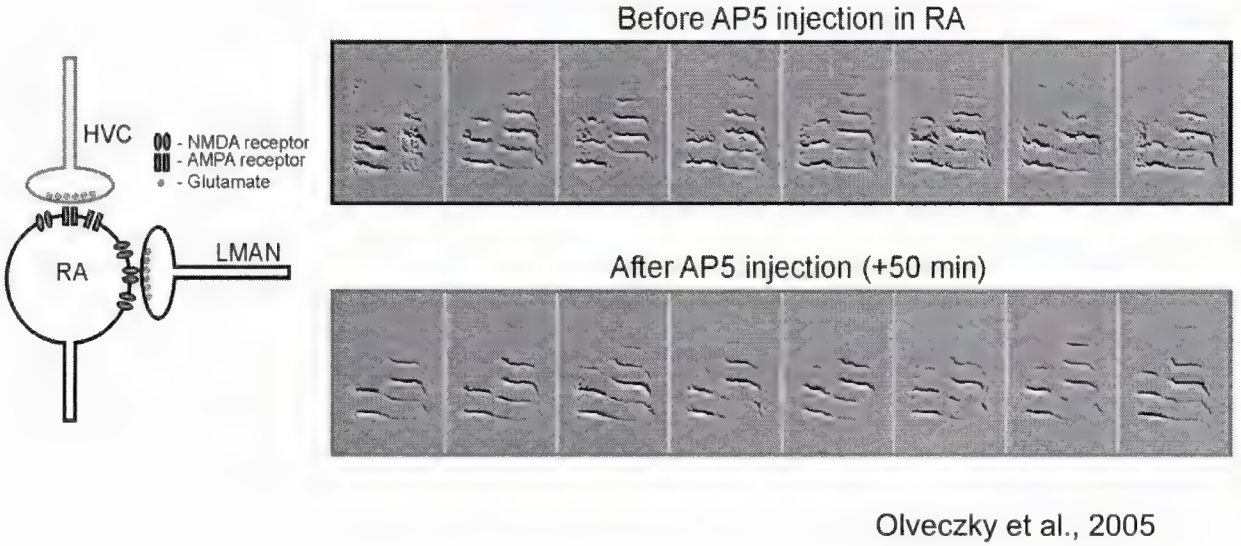






Kao et al., 2005

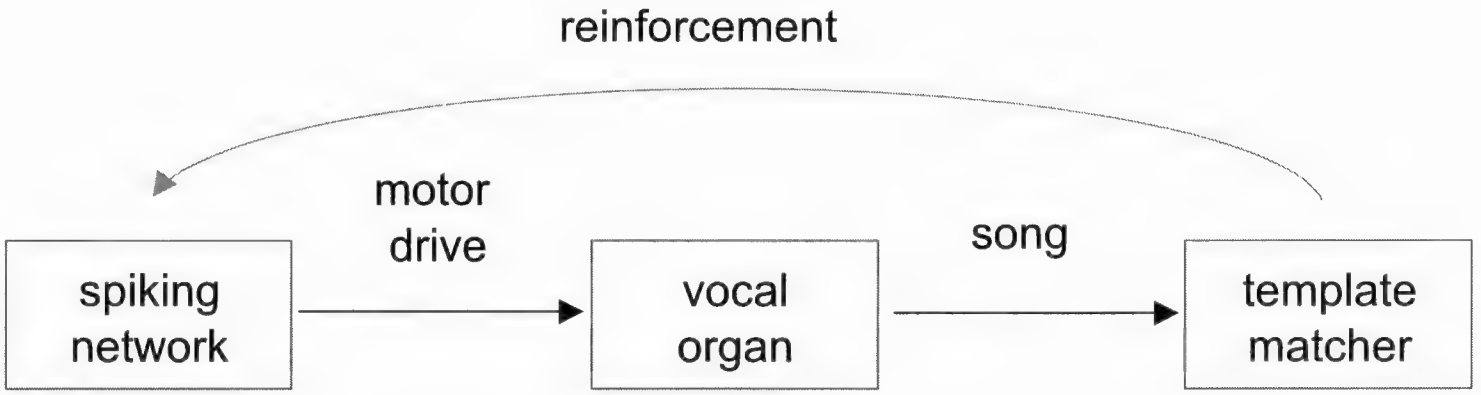
LMAN drives song variability



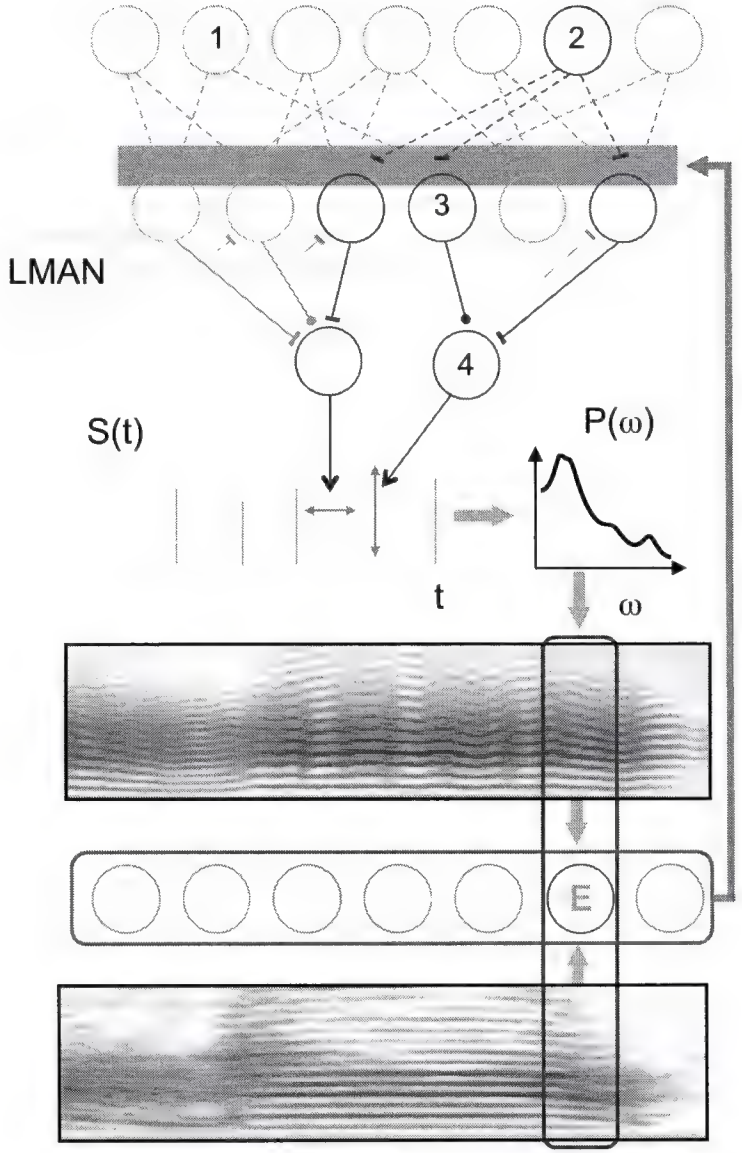
Olveczky et al., 2005

... through NMDA conductances (fast synaptic transmission)

Model of birdsong learning



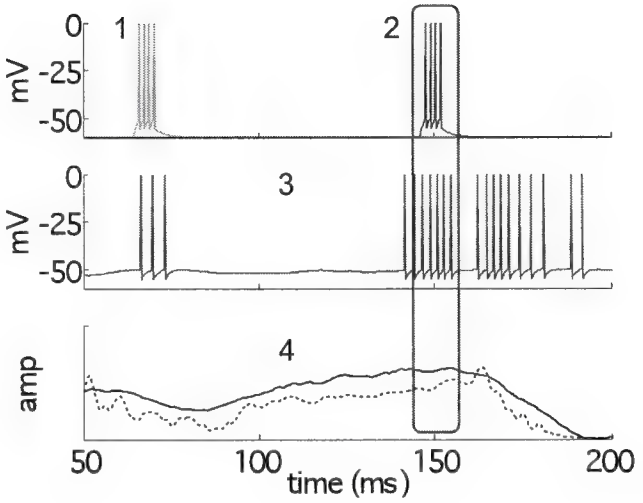
Fiete, Fee, Seung (2007)

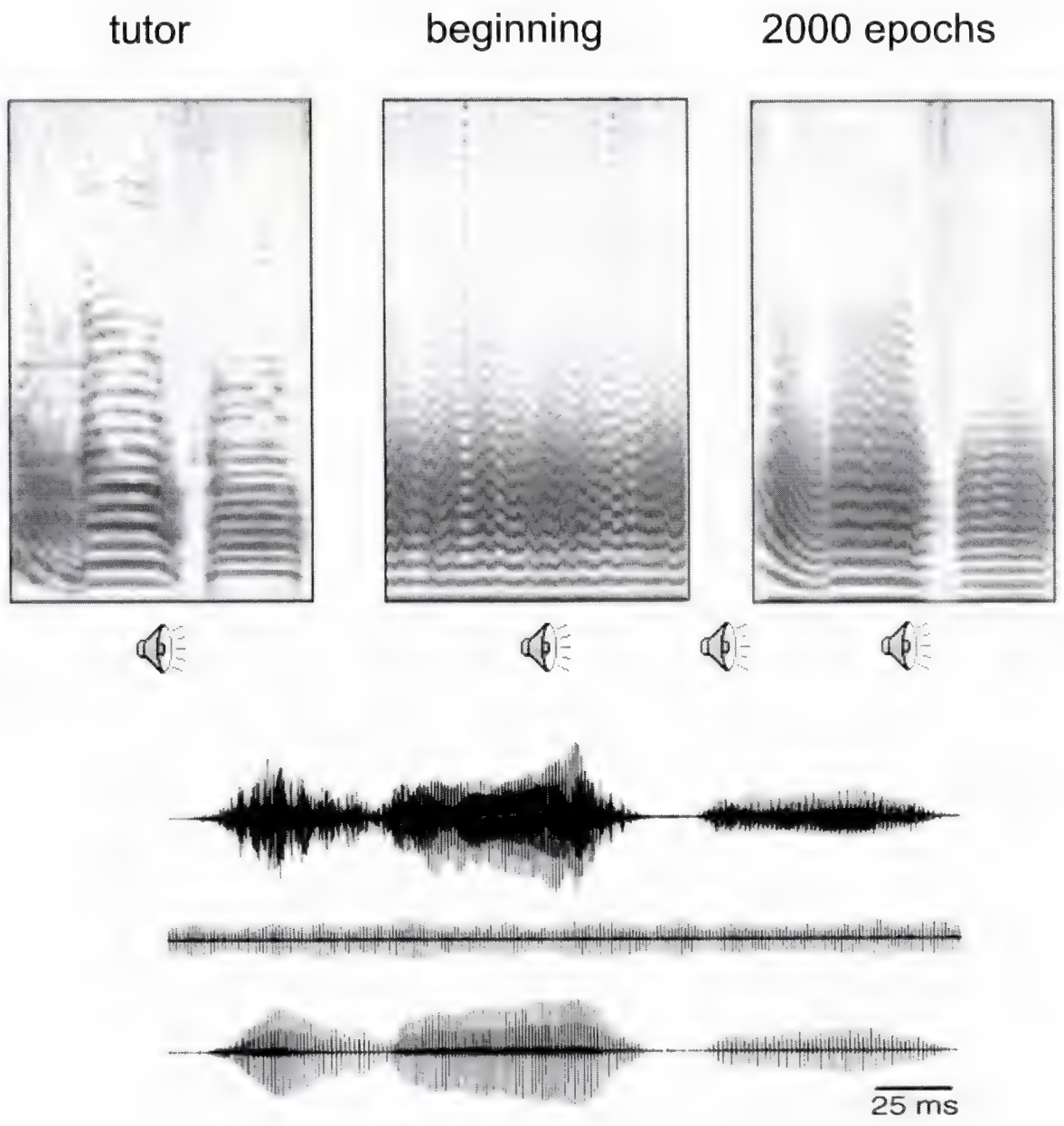


HVC

RA

motor



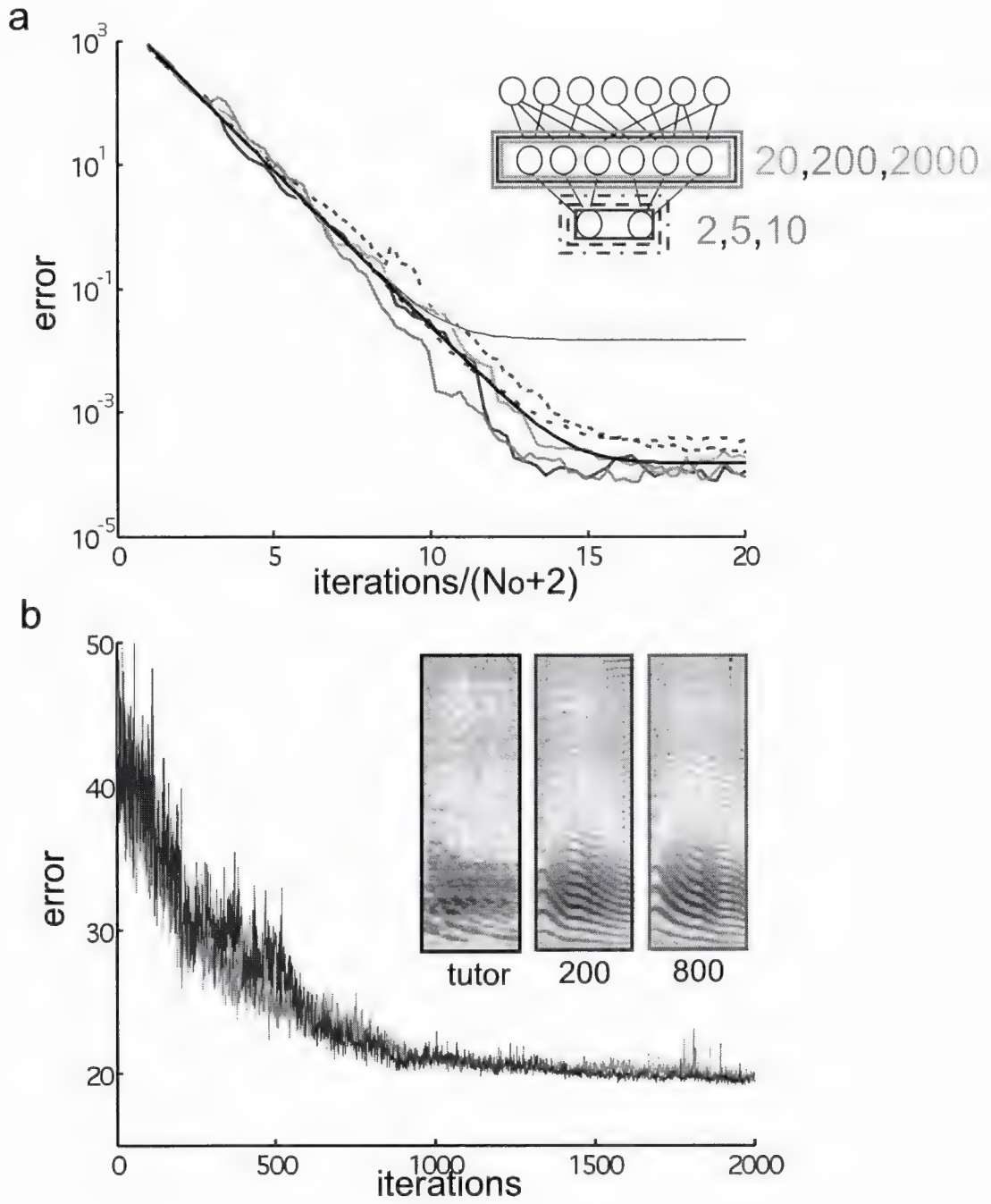

tutor
beginning
2000 epochs
25 ms

Scaling to large networks

Correlation of single neuron with reward = 1/N.
Time to learn ~ N?

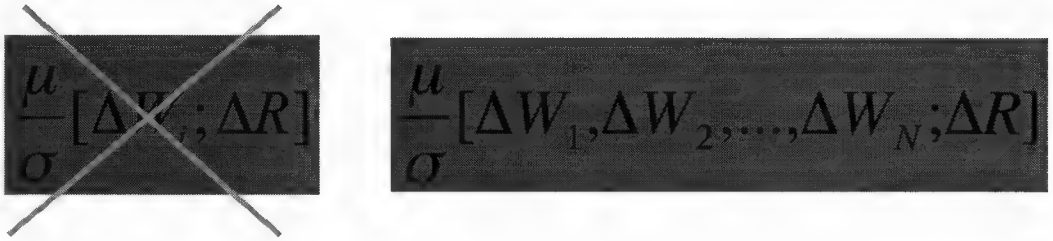
10-100 more neurons, 500-5000 more synapses in bird

Can learning by individual neural perturbation and reinforcement be fast enough to be useful in biologically large networks?


a
10^3
10^1
10^{-1}
10^{-3}
10^{-5}
error
0
5
10
15
20
iterations/(No+2)
20,200,2000
2,5,10
b
50
40
30
20
error
0
500
1000
1500
2000
iterations
tutor
200
800

Why does it work?

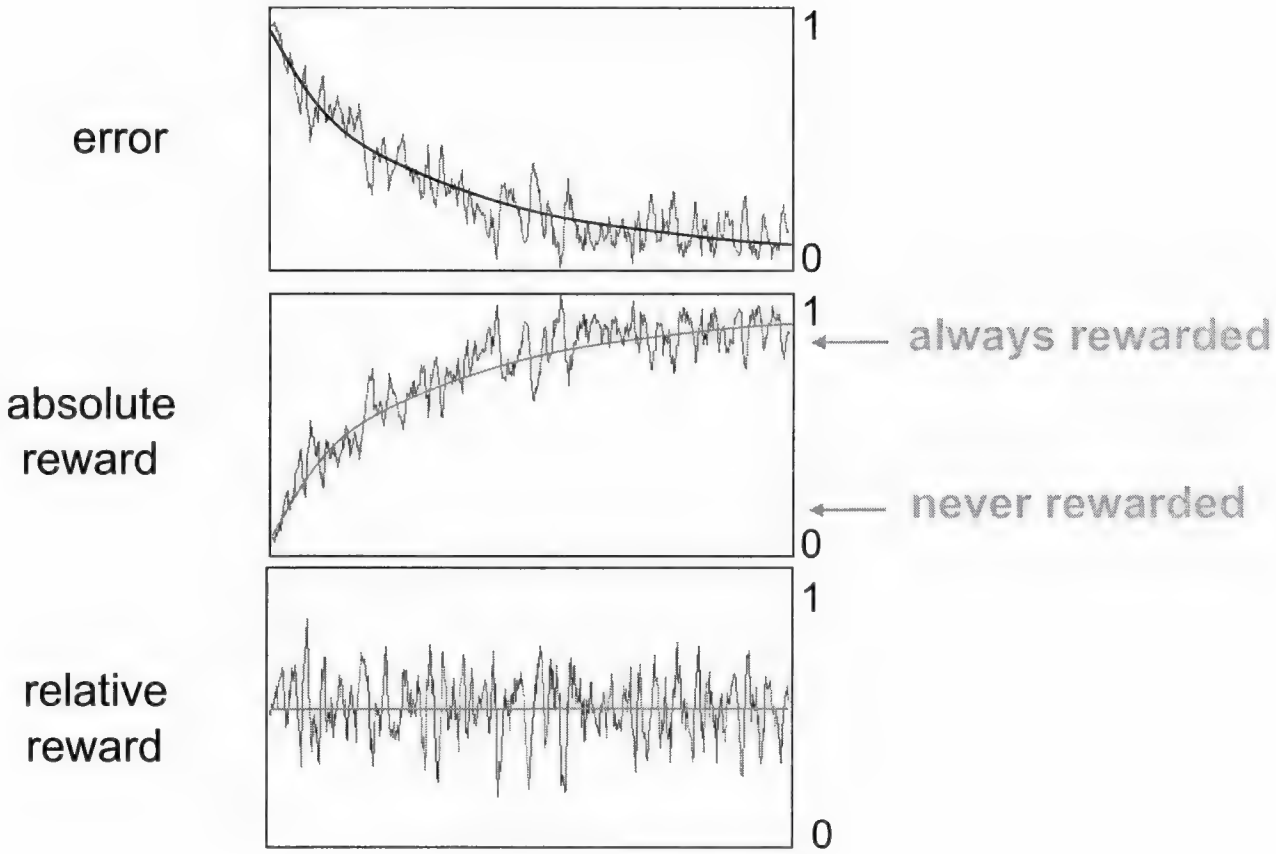
Redundant degrees of freedom in network



Individual correlations small,
but correlation of *summed* perturbation with reinforcement large.

Scales linearly with number of output degrees of freedom

Does reward increase as song improves?

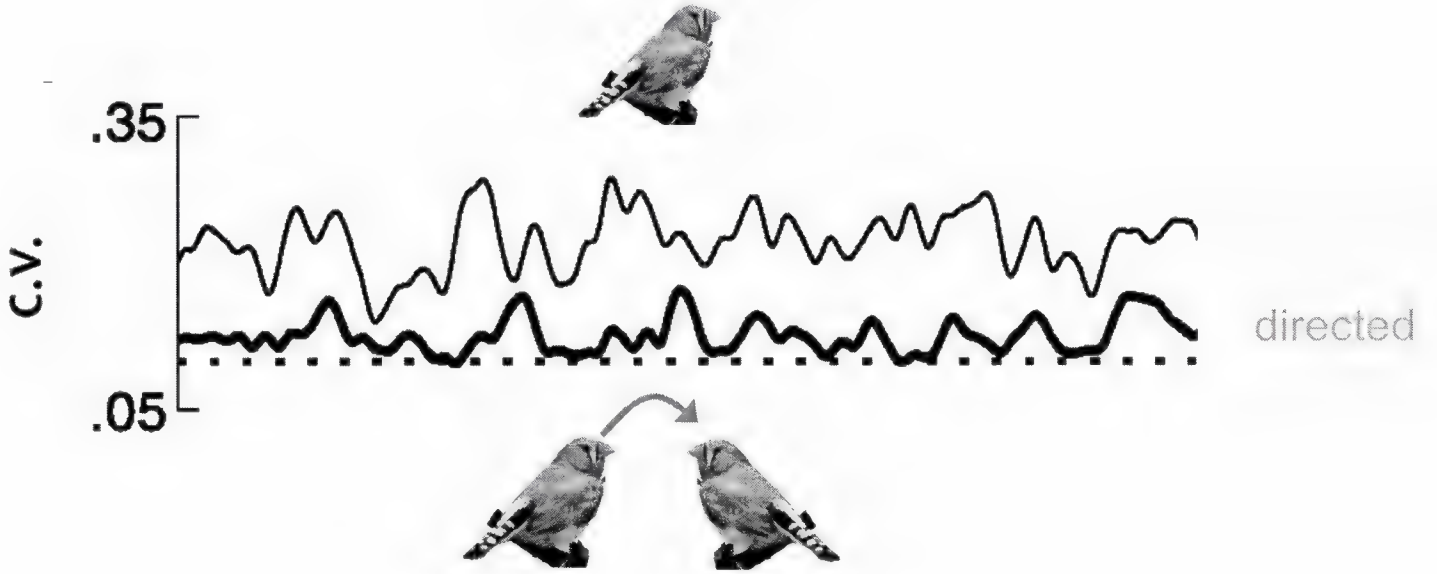


Differential, not absolute score important for learning

LMAN: Experimenter or supervisor?

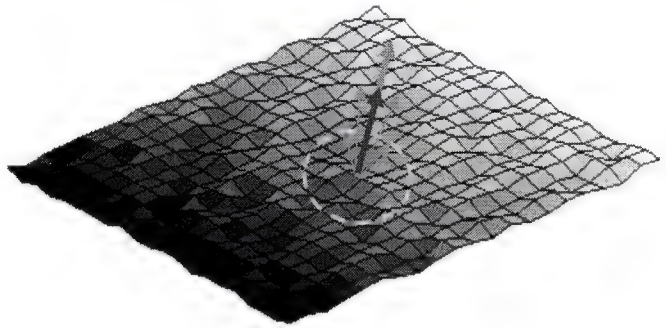
noise spike → reward → increase active weights
no noise spike → reward → decrease active weights

External experimenter: smart perturbations



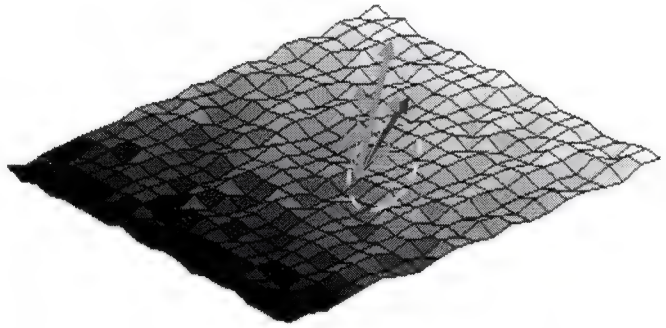
Hessler and Doupe, 1999

Faster reinforcement learning in realistic networks



uniform exploration,
selective reinforcement
→ uphill

- how might brain learn statistics of exploration?



biased exploration,
selective reinforcement:
→ uphill?

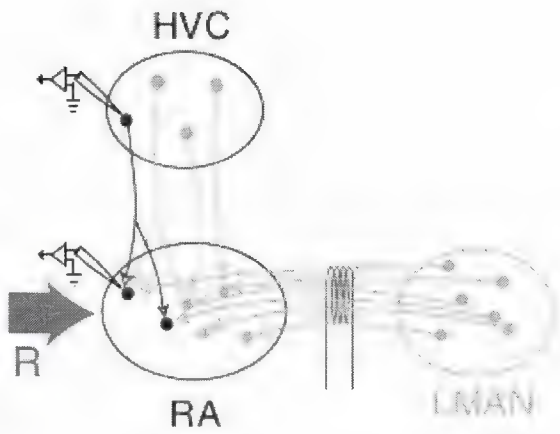
- how does brain reduce anisotropy using only local information?
- how does brain form, train modular networks (hierarchy)?

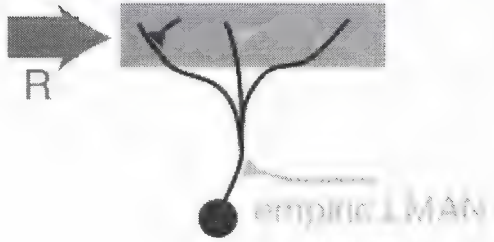
‘Simple’ tasks can be learned

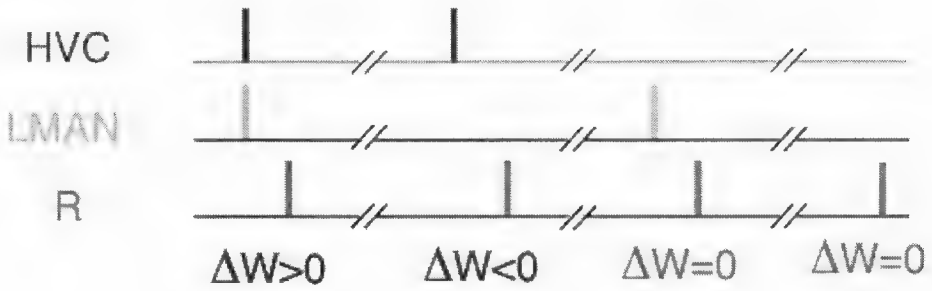
- Learning time independent of network size in redundant networks.
- Learning by single-neuron experimentation fast enough to be consistent with fine motor skill learning: e.g., zebra finches practice ~100,000 times.

Trial-and-error motor learning may be driven by experimentation and reinforcement in individual neurons

Experiment





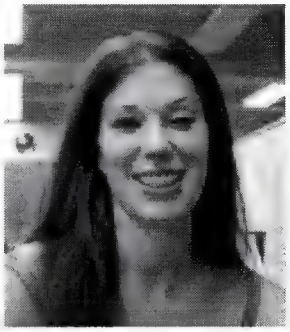


The multiple timescales of motor memory

Reza Shadmehr
Johns Hopkins School of Medicine

Jun Izawa

Haiyin Chen-Harris



Sarah Hemminger

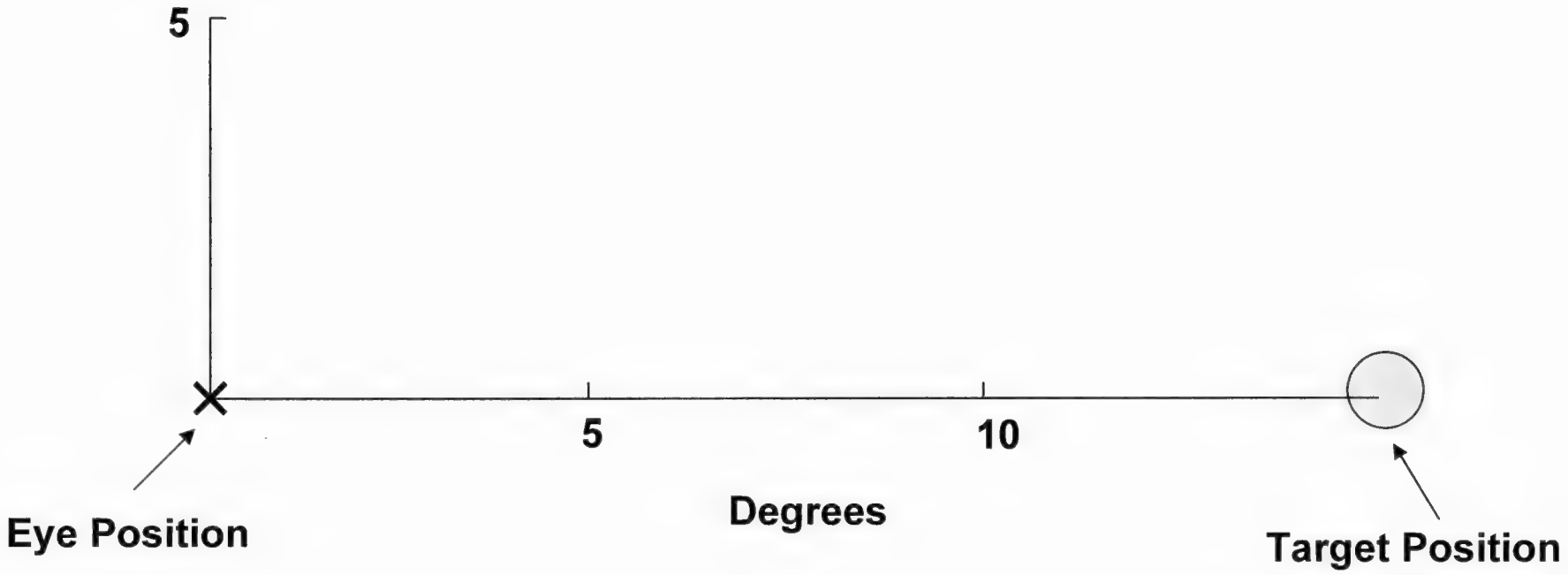
Minnan Xu

Dave Zee

Two interesting observations in memory research

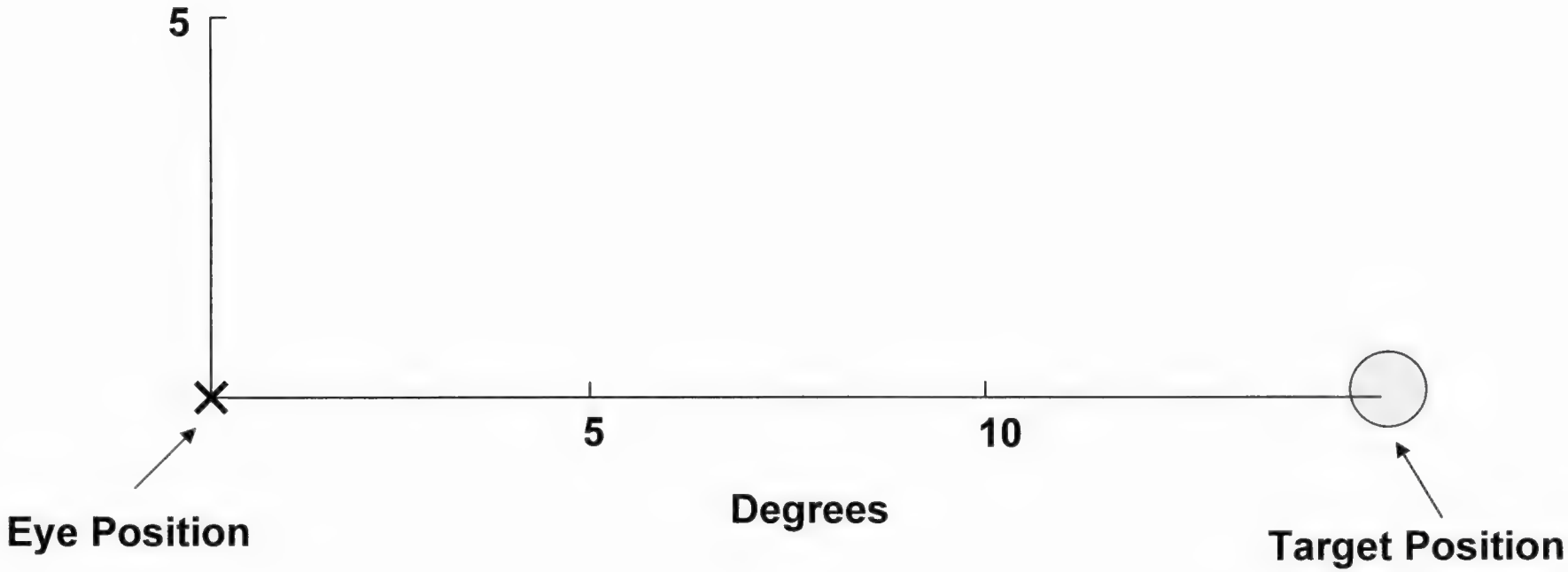
1. Savings of memory
   "de-adaptation" does not seem to erase the memory that was produced by adaptation.
2. Spontaneous recovery of memory
   Passage of time seems to alter the contents of memory.

Saccade adaptation: gain decrease



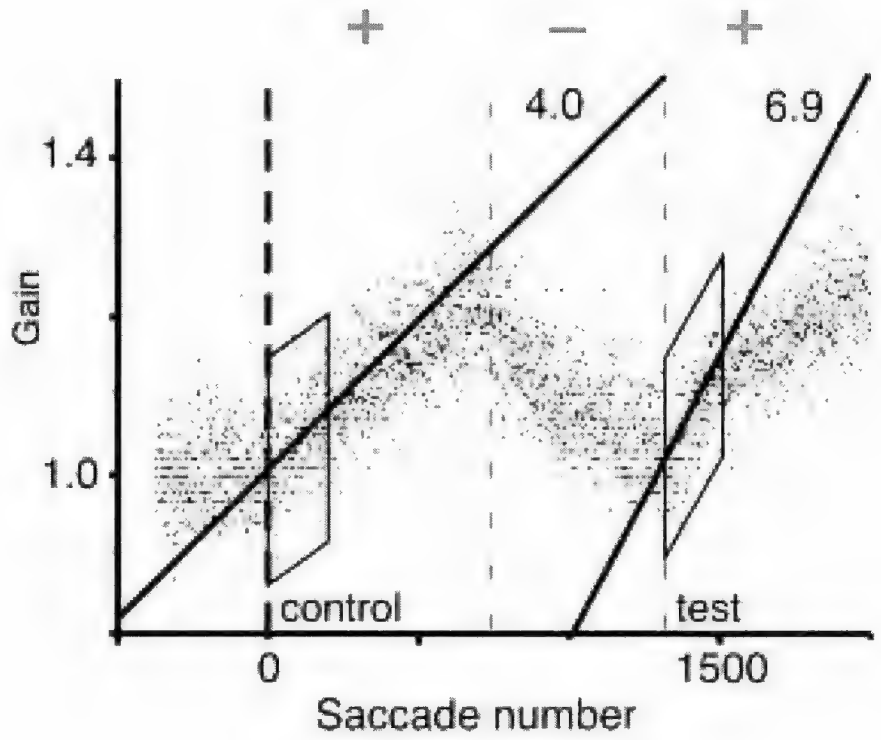
McLaughlin 1967

Saccade adaptation: gain decrease



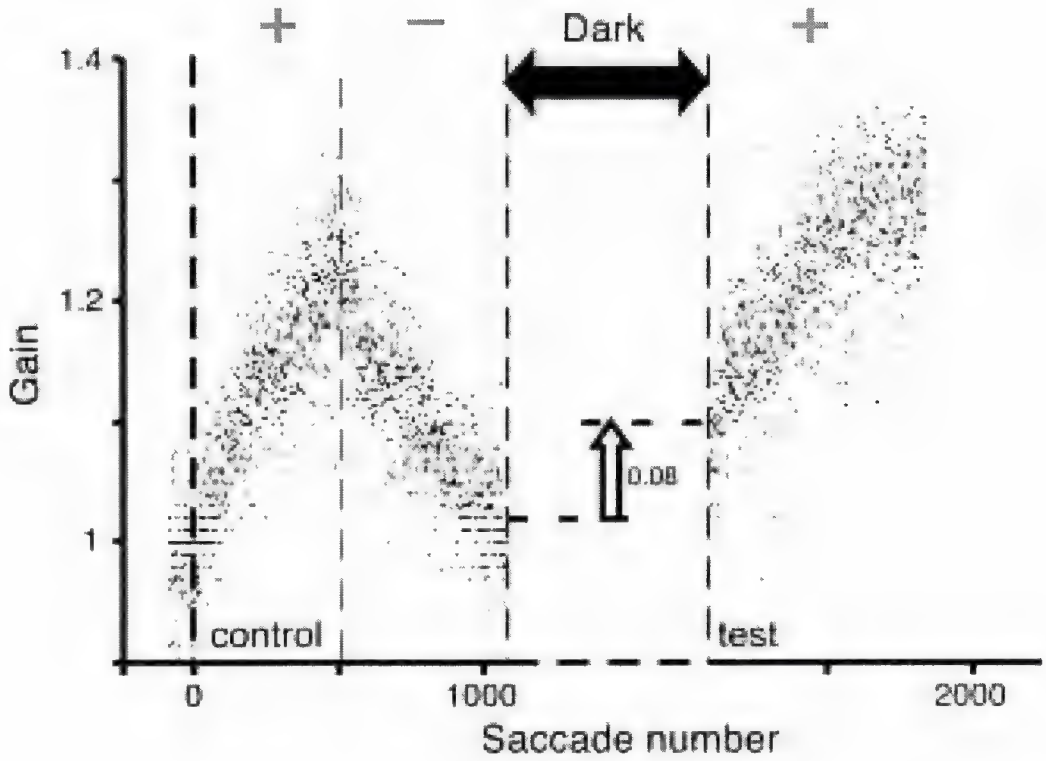
McLaughlin 1967

Savings: when adaptation is followed by de-adaptation, motor system still exhibits recall



Yoshiko Kojima, Yoshiki Iwamoto, and Kaoru Yoshida (2004) J Neurosci 24:7531.

The multiple timescales of motor memory



Yoshiko Kojima et al. (2004) J Neurosci 24:7531.

The multiple timescales of motor memory

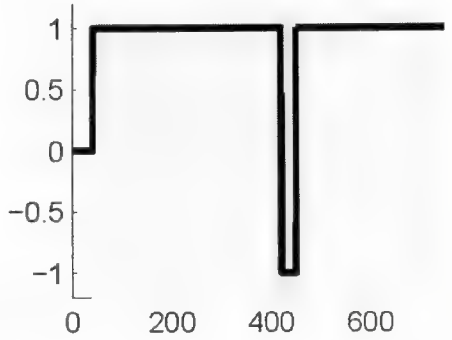
A fast learning system that forgets quickly

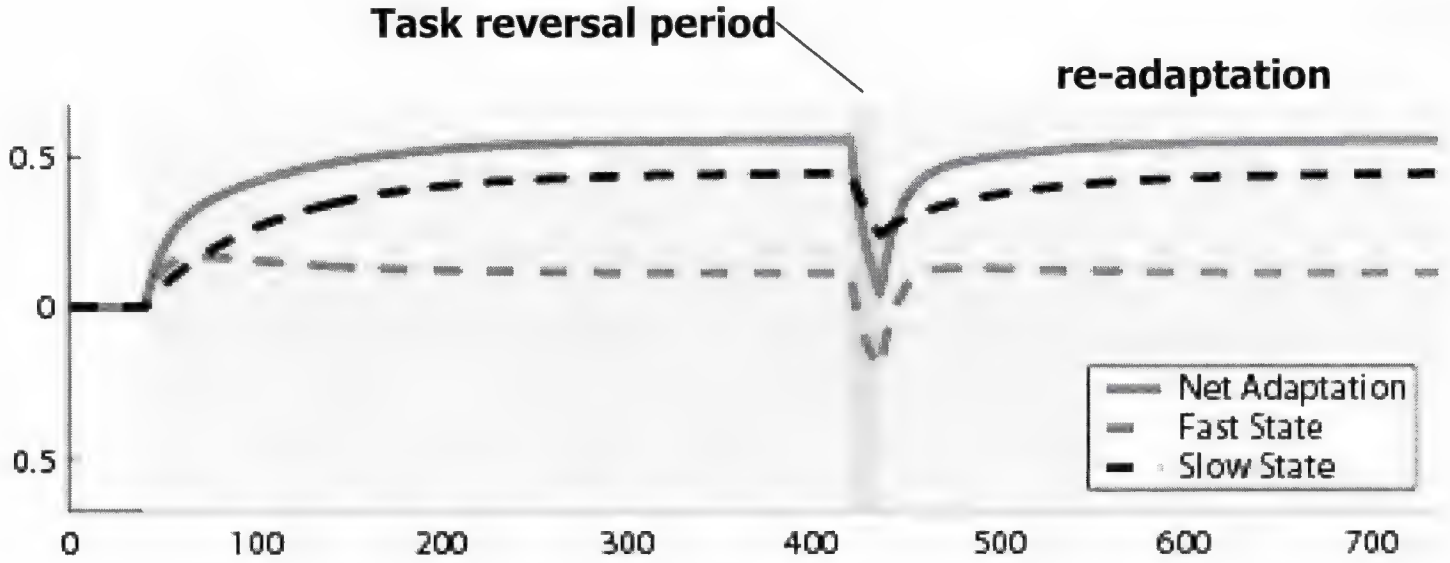
A slow learning system that hardly forgets

prediction $$\hat{y}^{(n)} = \hat{y}_f^{(n)} + \hat{y}_s^{(n)}$$

Smith et al. (PLOS Biology, 2006)

Savings: de-adaptation may not erase adaptation

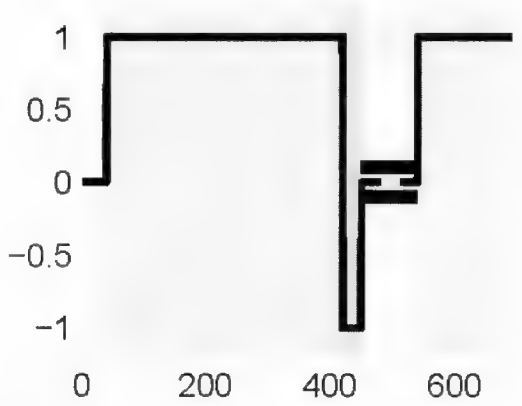


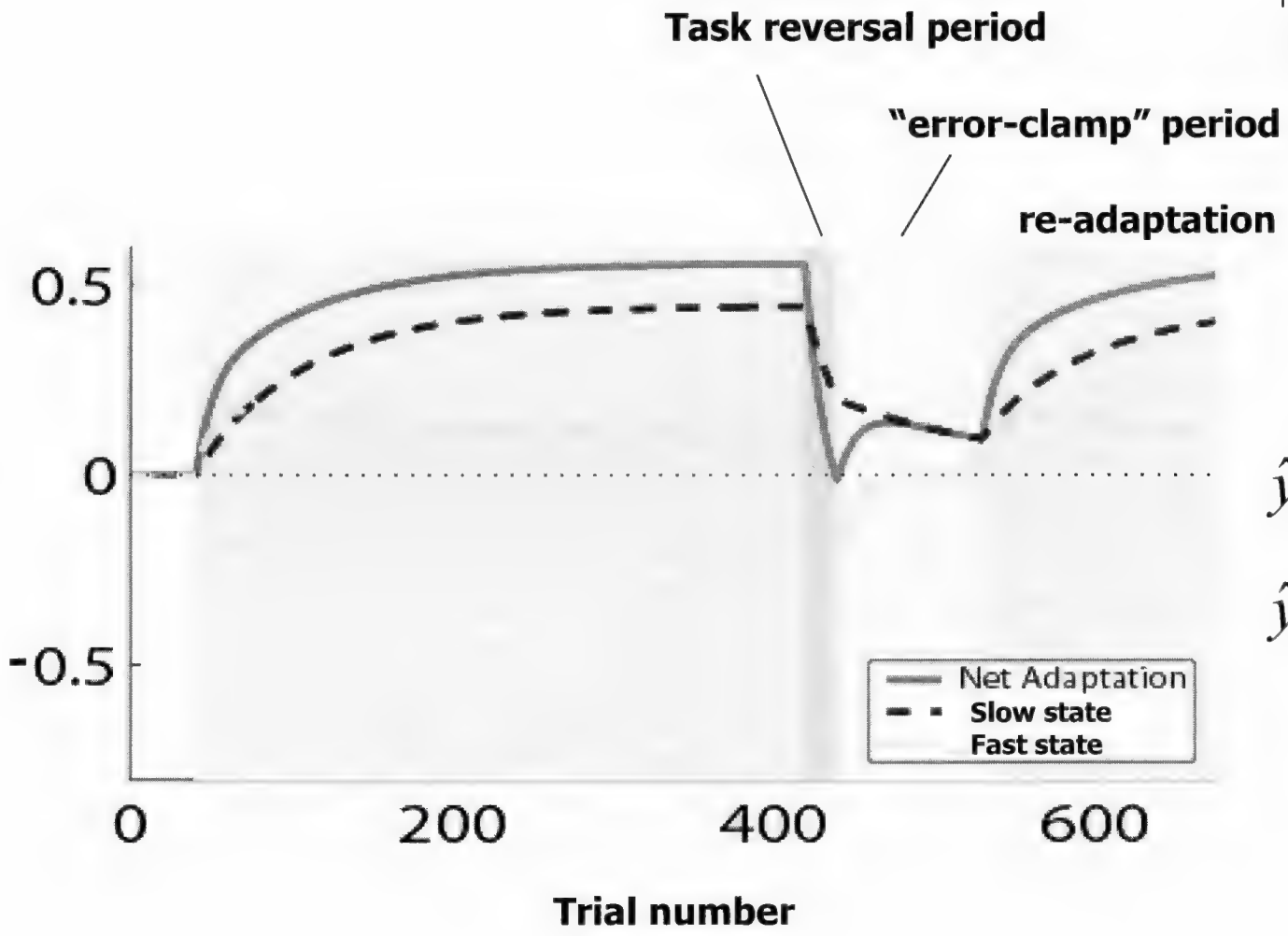


Trial number

Smith et al. PLOS Biology, 2006

Spontaneous recovery: Passage of time has asymmetric affects on the fast and slow systems



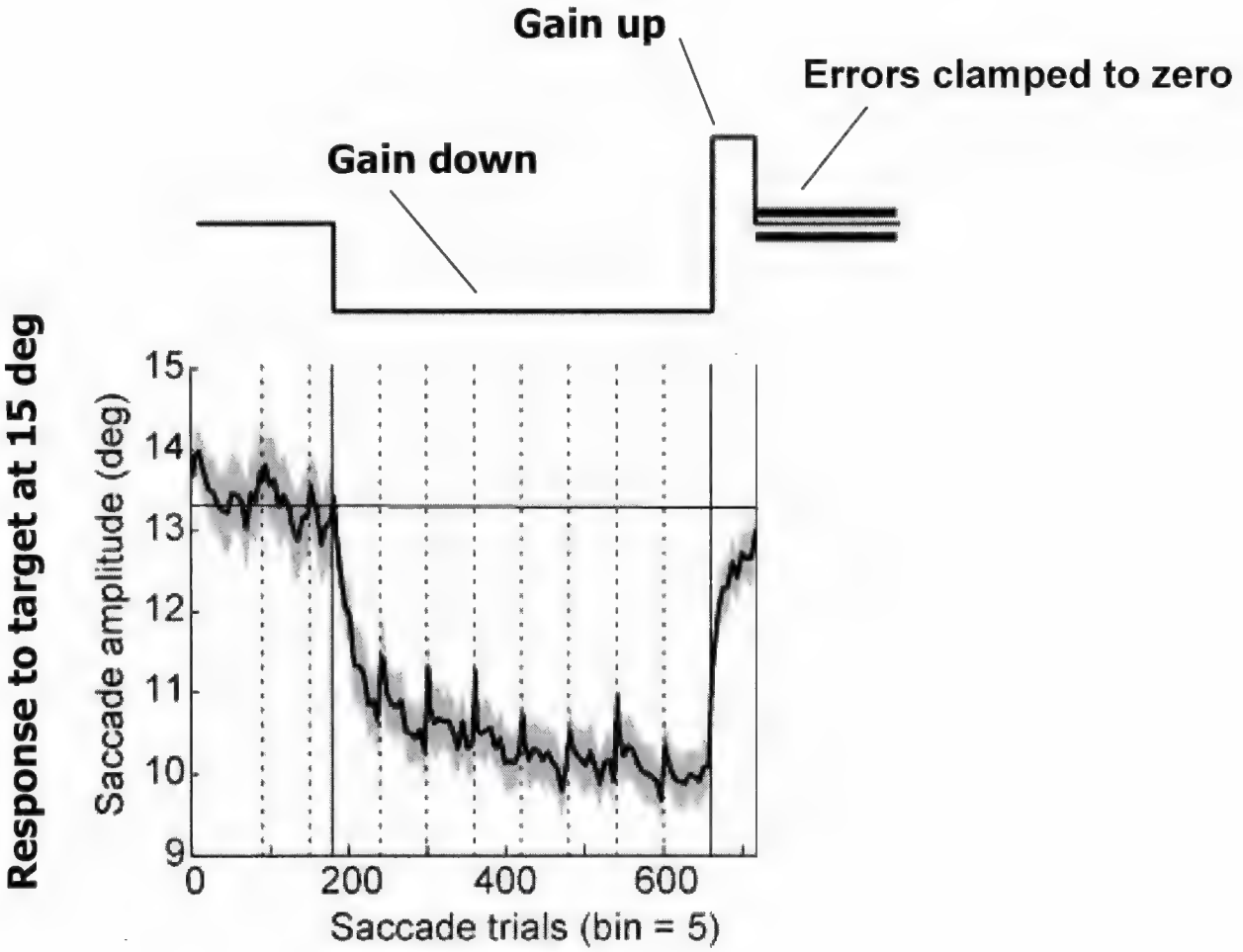


$$\hat{y}_1^{(n+1)} = a_1 \hat{y}_1^{(n)} + b_1 \tilde{y}^{(n)}$$

$$\hat{y}_2^{(n+1)} = a_2 \hat{y}_2^{(n)} + b_2 \tilde{y}^{(n)}$$

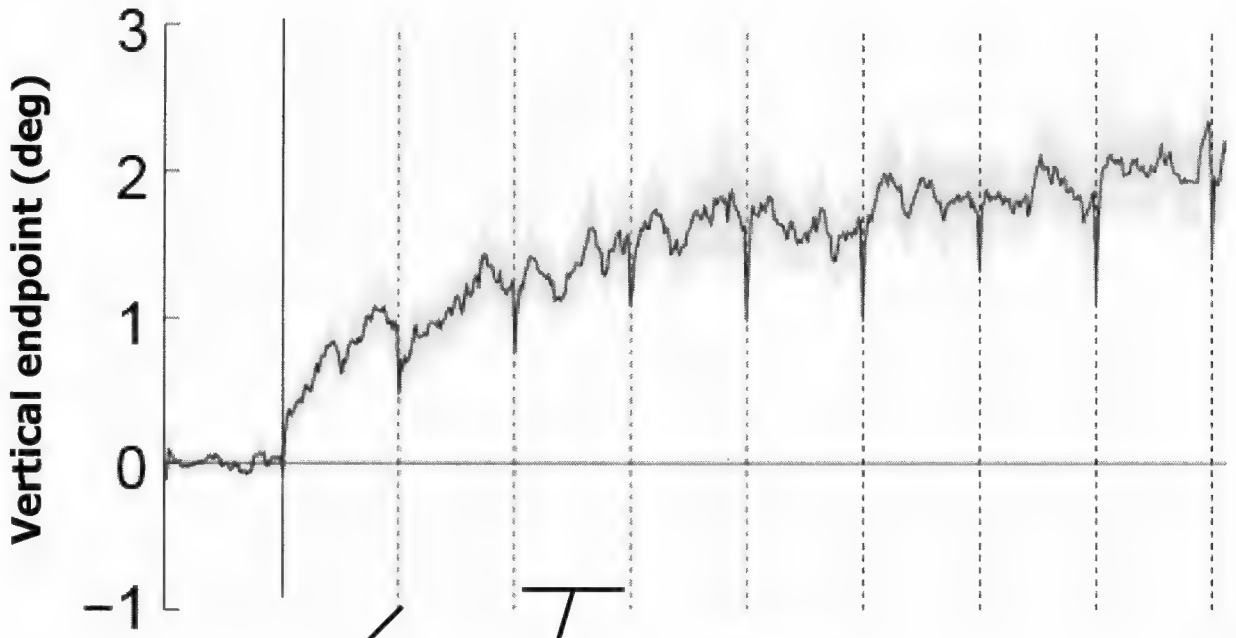
Smith et al. PLOS Biology, 2006

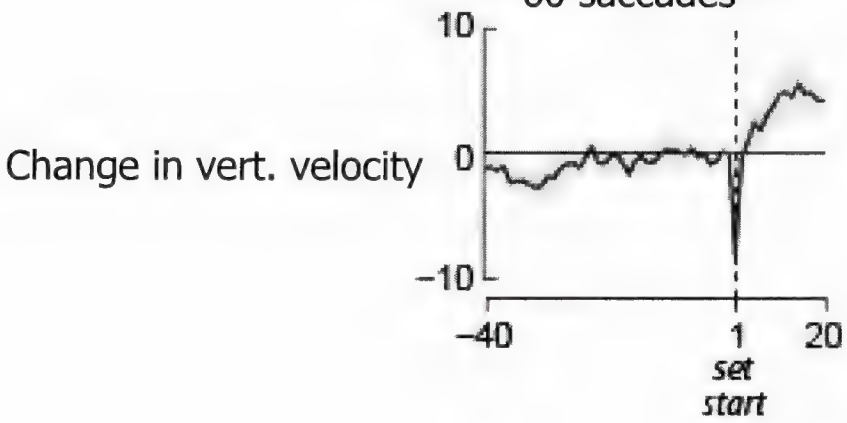
Spontaneous recovery in saccade adaptation

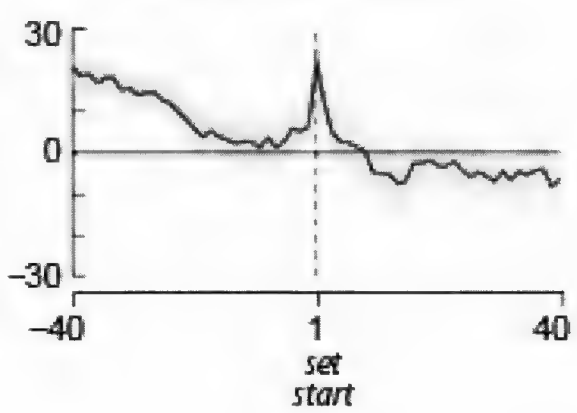


Ethier et al. (J Neurophysiol, 2008)

'Forgetting' as a signature of a fast adaptive system

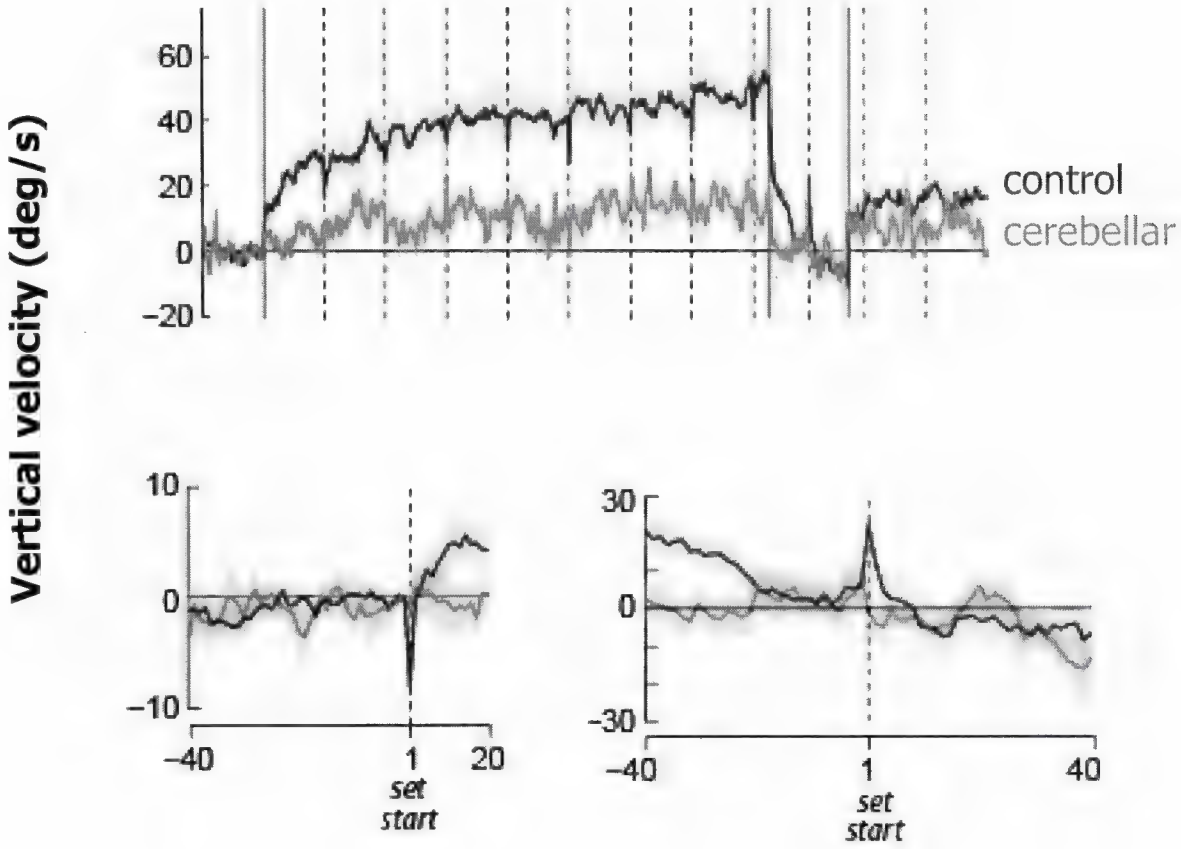






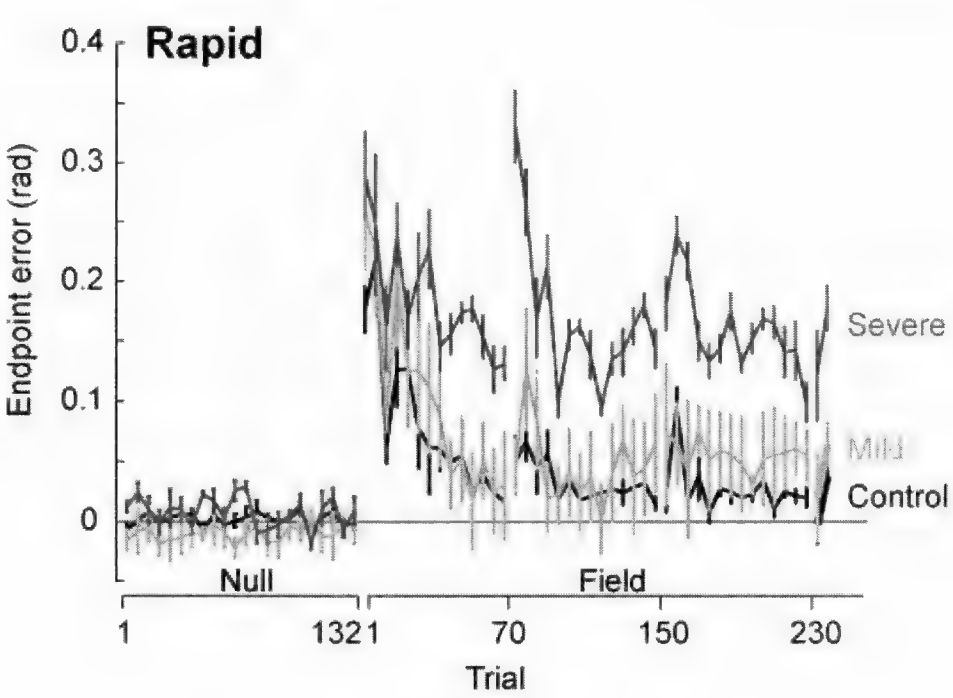
Minnan Xu, unpublished data

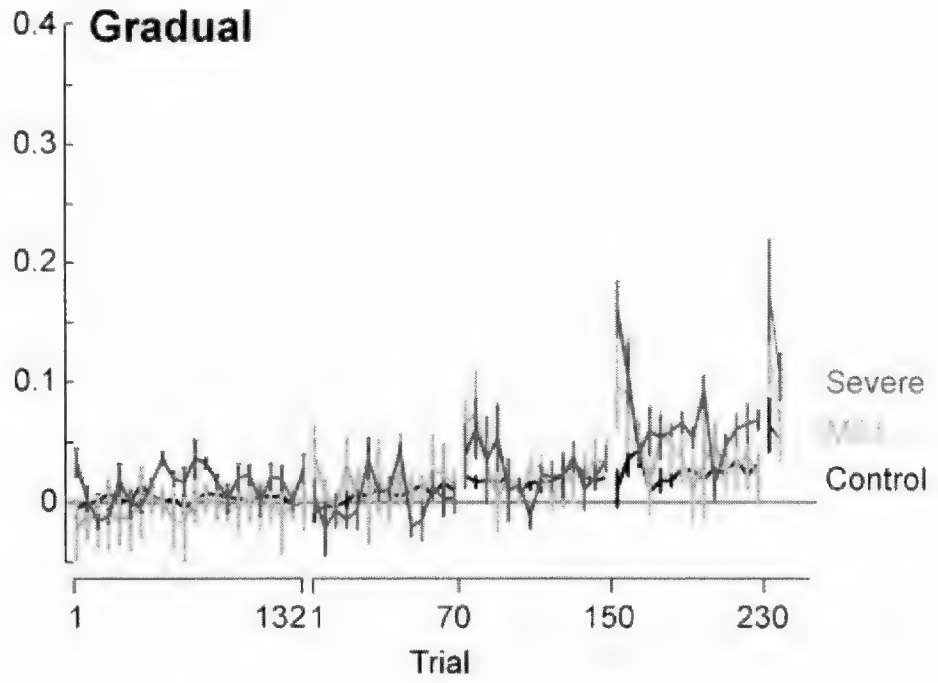
Cerebellar cortical damage appears to affect predominately the fast timescale of adaptation

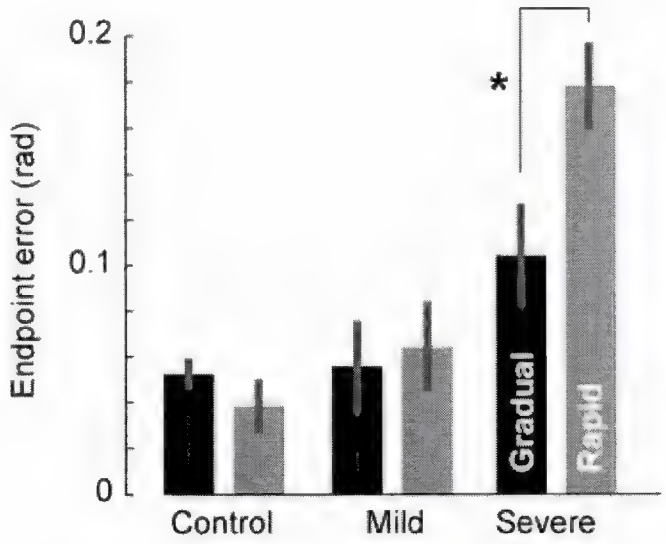


Minnan Xu, unpublished data

Despite severe cerebellar degeneration, improved learning if perturbations are imposed gradually



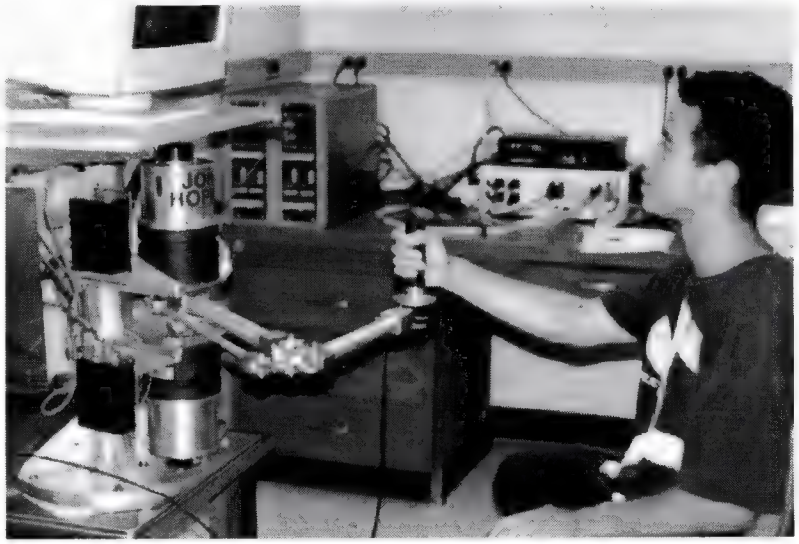


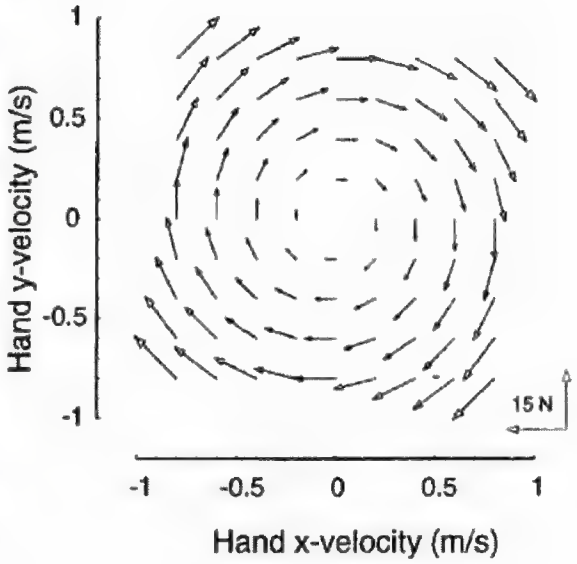


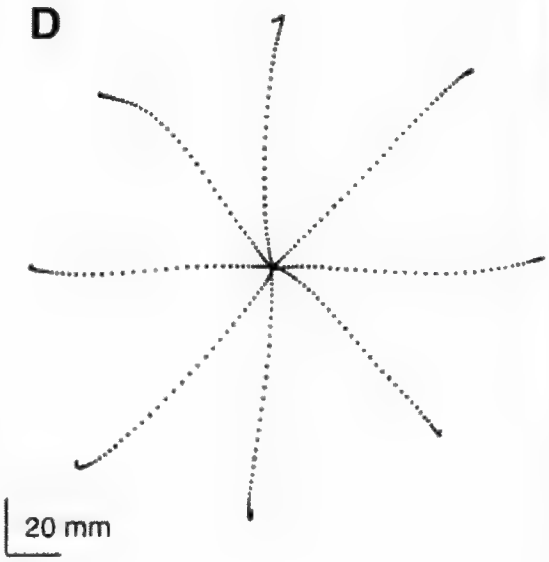
Sarah Hemminger, unpublished data

Spontaneous recovery in reach adaptation

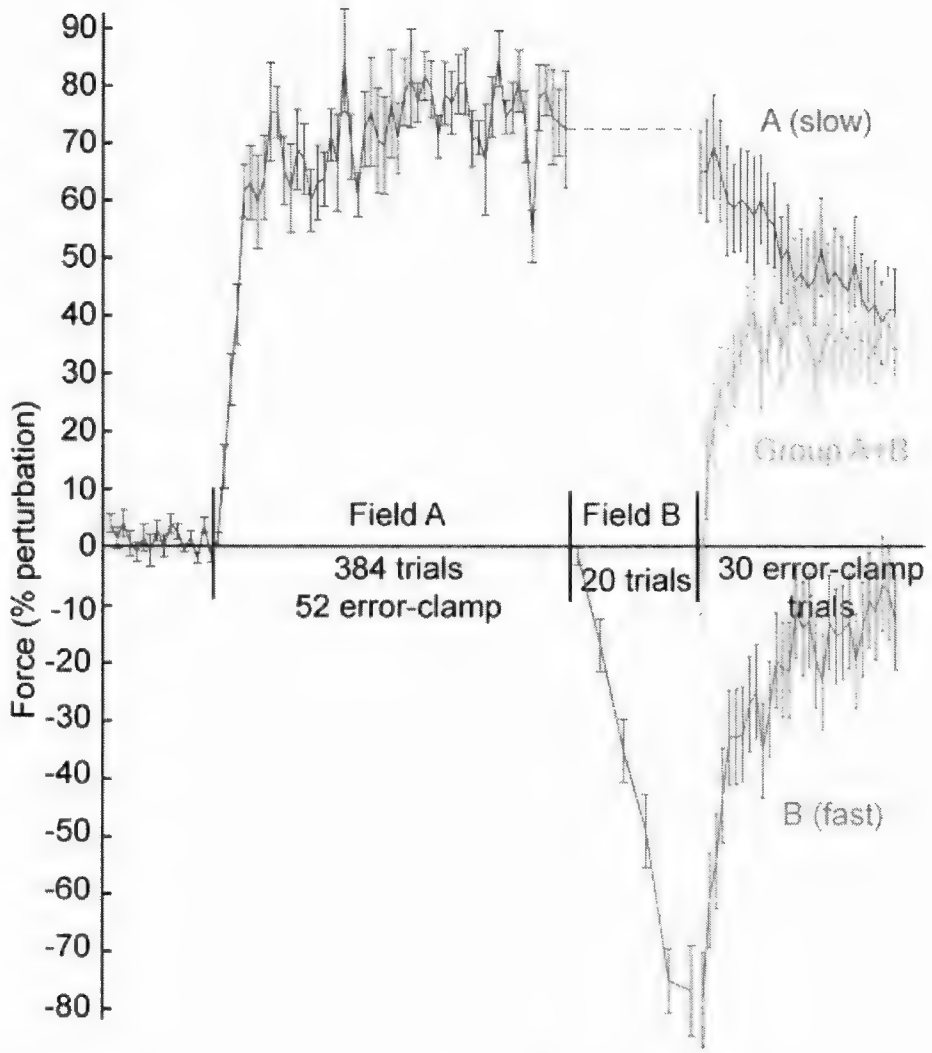
Two groups of subjects: **group A**
group A+B







Hemminger and Shadmehr, J Neurosci (2008)



Learning of field B did not erase A, but produced a memory of B that competed with A. Memory of B was a fast memory: within 30 trials, it washed away, producing spontaneous recovery of A.

Hemminger and Shadmehr, J Neurosci (2008)

Summary

Maurice Smith

A prediction error causes changes in multiple adaptive processes. Some are highly responsive to error, but rapidly forget. Others are poorly responsive to error but have high retention.

Minnan Xu

In cerebellar damage, the fast timescale of adaptation appears to be missing but the slow timescale may be less affected.

Sarah Hemminger

There may be no such thing as unlearning. Unlearning may be simply instantiation of a new learning process that competes with previous learning.

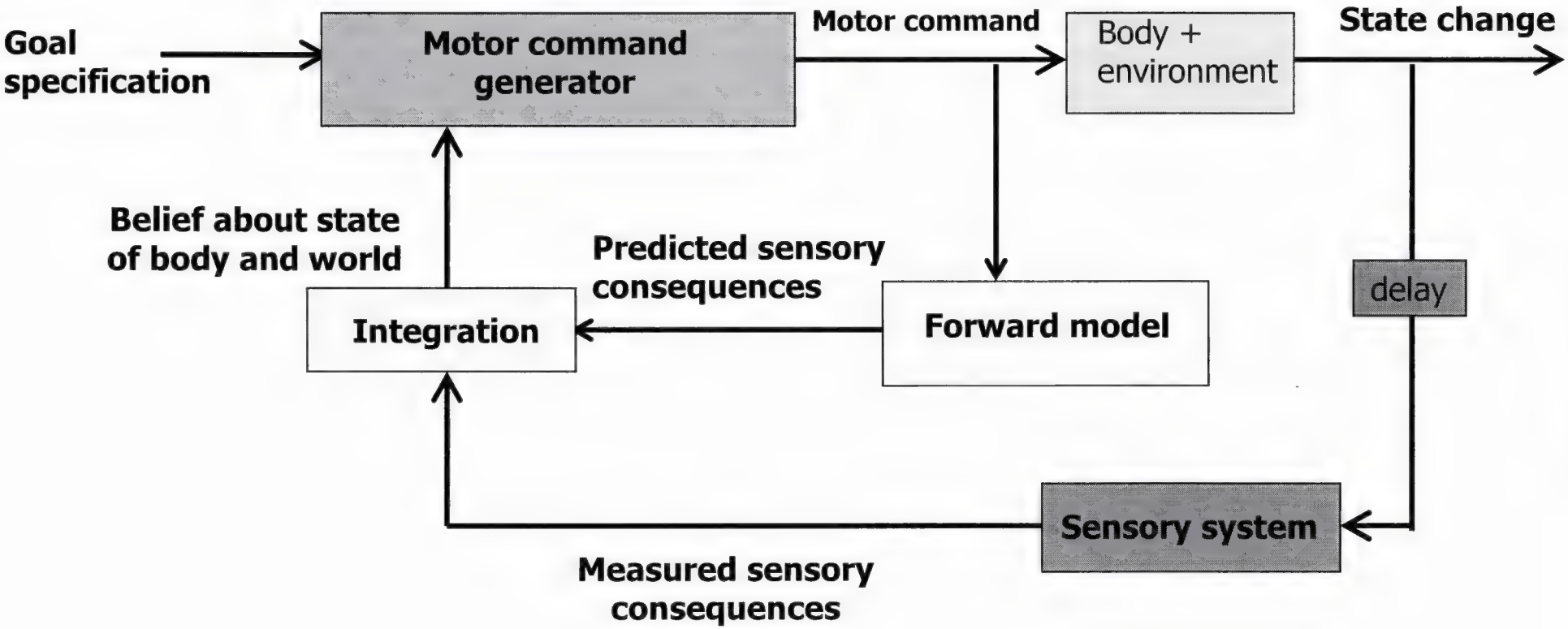
What are some of the holes in this idea?

1. Movements have trajectories and temporal properties. Our model so far is a cartoon: a movement as a single data point. What about a control system?

2. How do the fast and slow timescales of learning fit into the control system?

PRINCIPLES OF
NEURAL SCIENCE

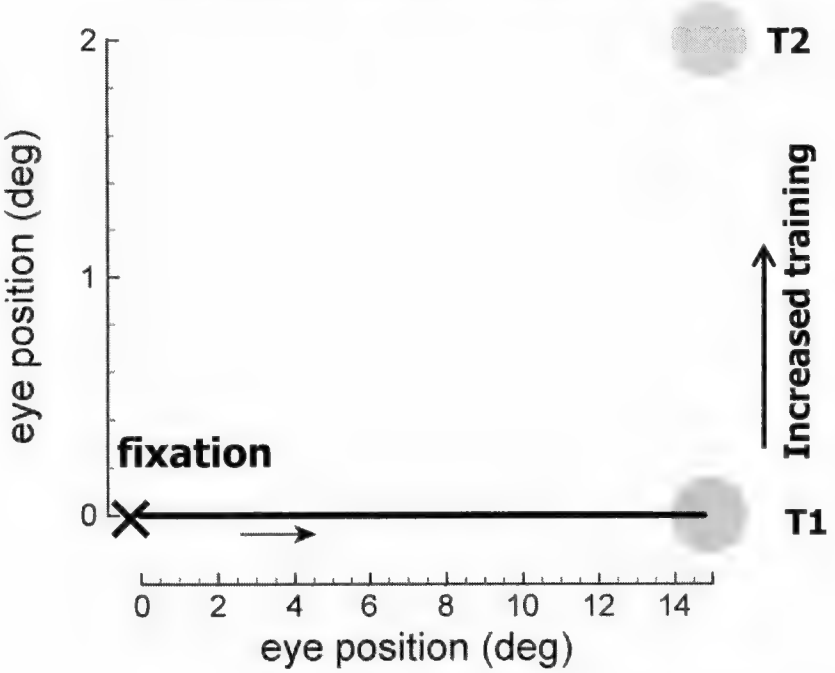
A control scheme for motor control



Cross-axis saccade adaptation

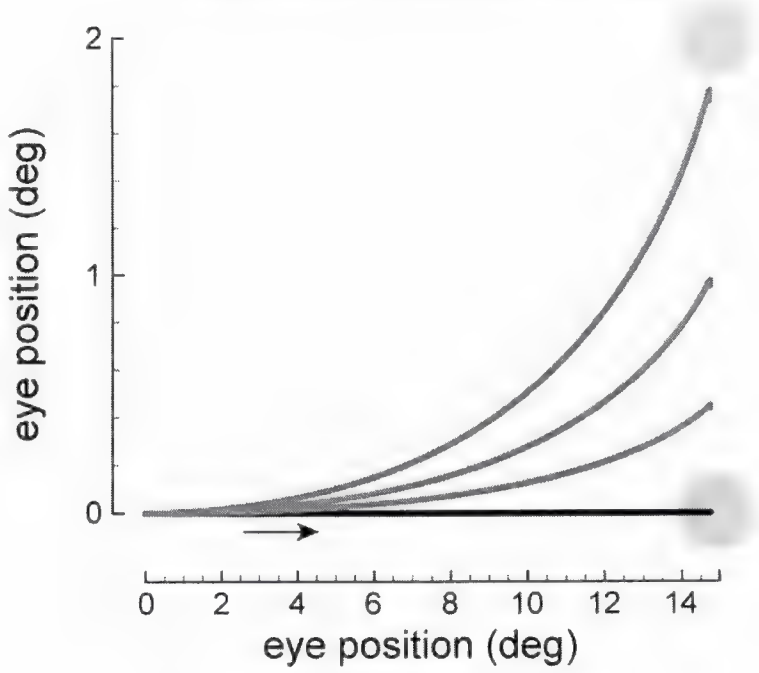
Equal rates of learning in the controller and the forward model

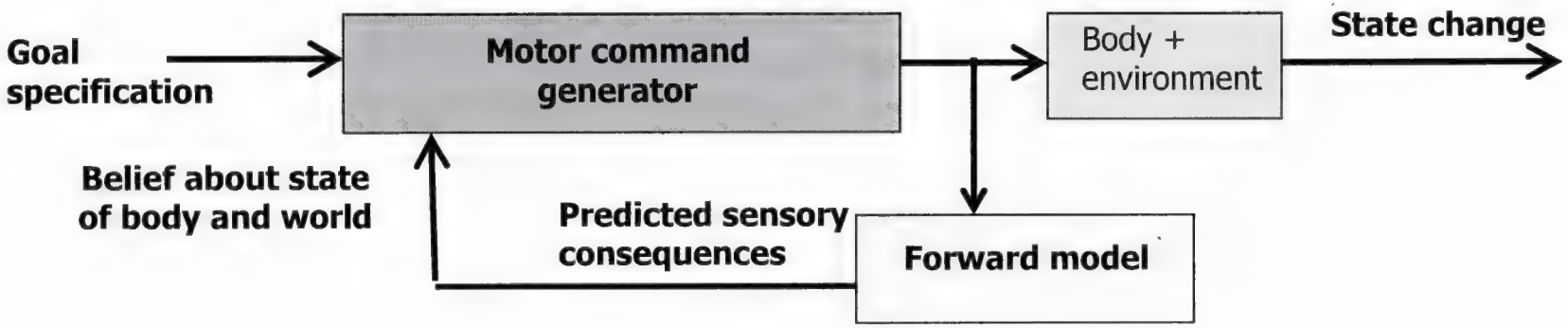
saccades remain straight



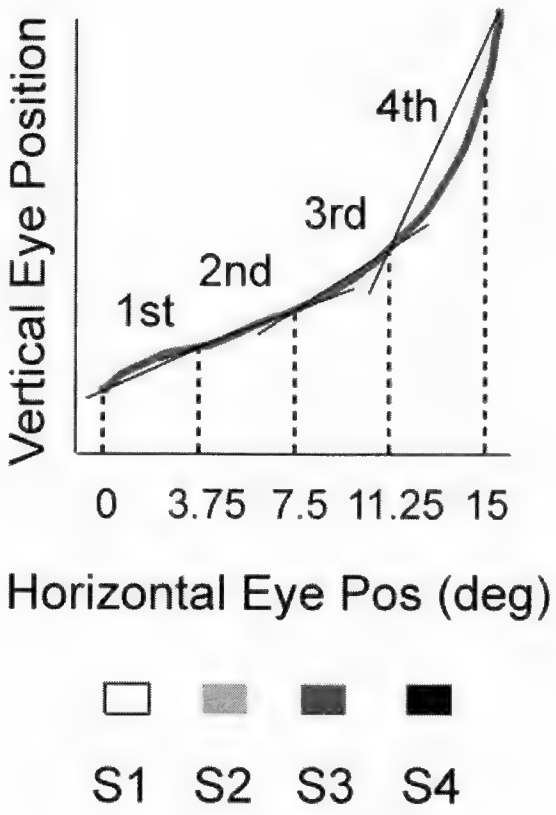
Learning in the forward model only

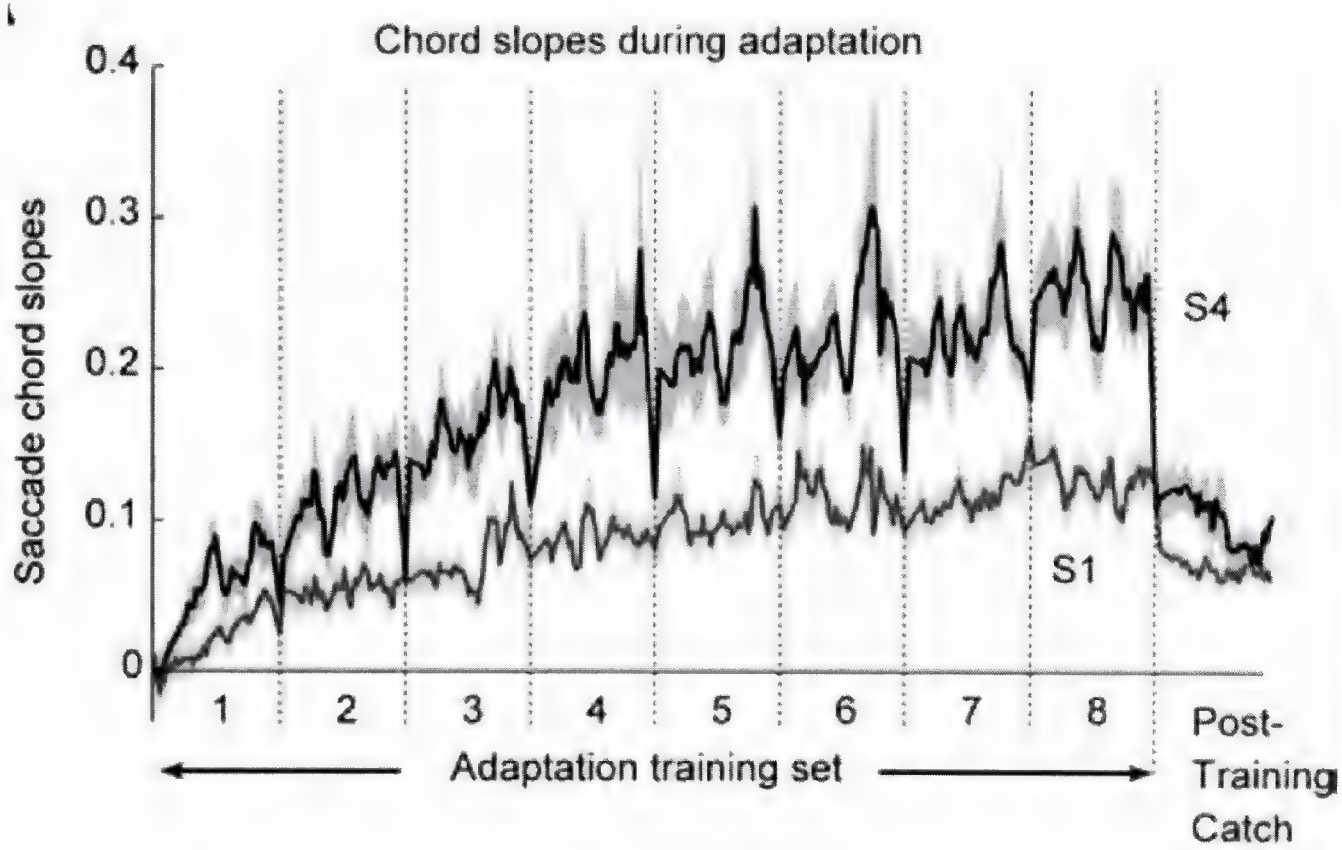
saccades become curved





Curvature of saccades quantified through chord slopes





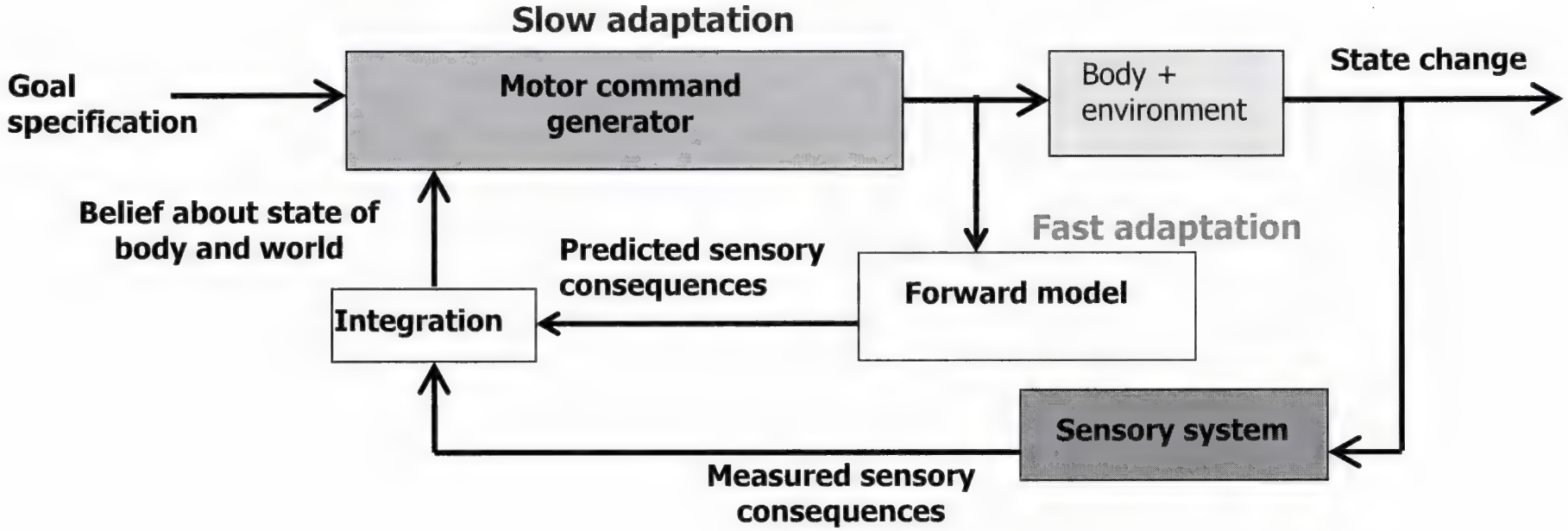
Chen-Harris et al. (J Neurosci, 2008)

Summary

Haiyin Chen-Harris

Motor commands that initiate a movement adapt more slowly than the commands that "steer" the movement.

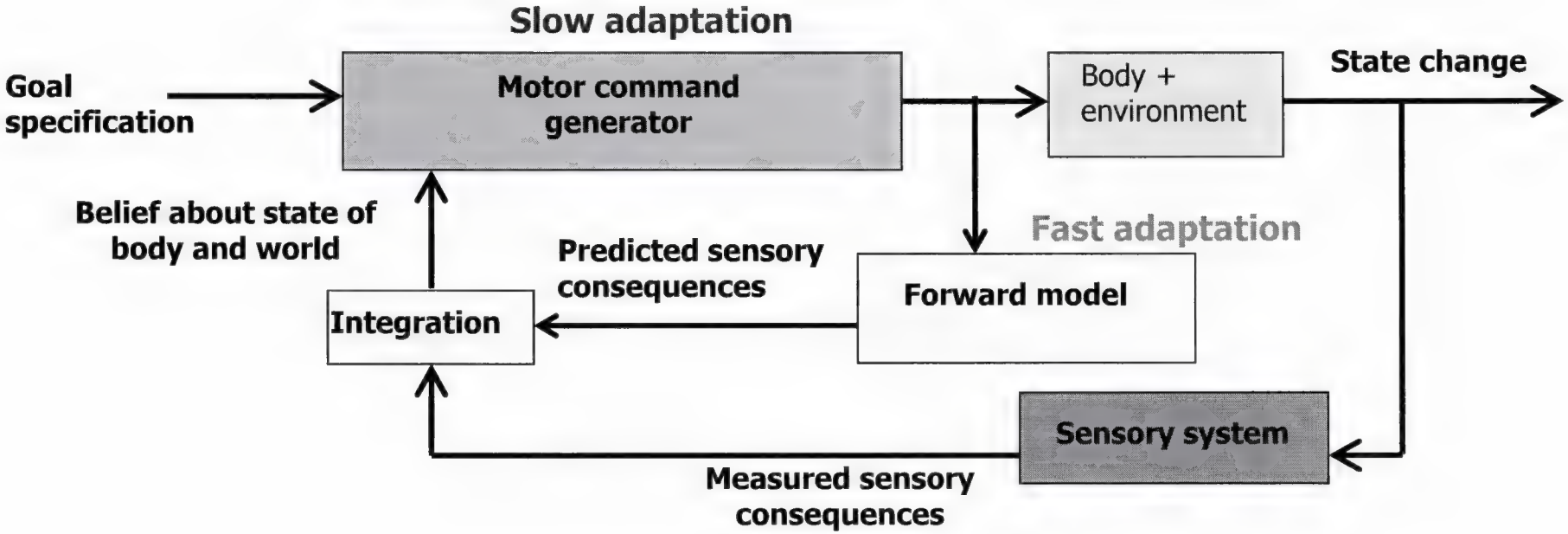
Perhaps the fast adaptive system is in the 'forward model'.



What are some of the holes in this idea?

1. What exactly do you we mean by a 'motor command generator'?

2. How do these 'boxes' relate to the brain?



Because of our wisdom, we will travel far for love,

As all movement is really a sign of thirst,

And speaking really says "I am hungry to know you."

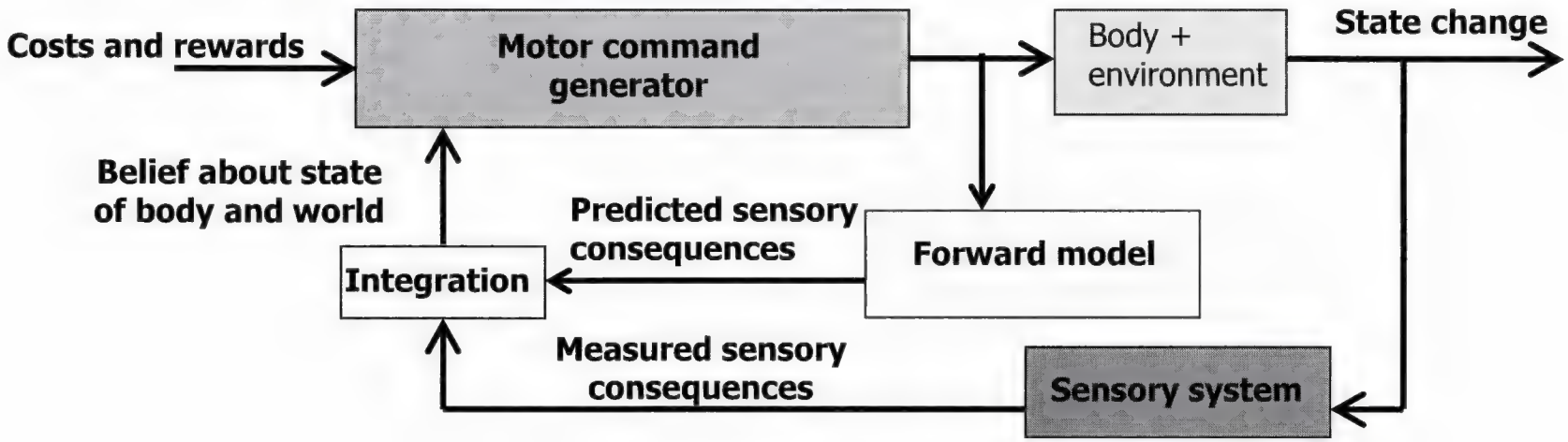
Hafez, 14th century Persian poet

The purpose of a movement is to achieve a more rewarding state

Motor control is best viewed as a problem in optimization (Todorov 2005)

Actual state of the system
(eye state, target state, etc.)

Signal dependent motor noise

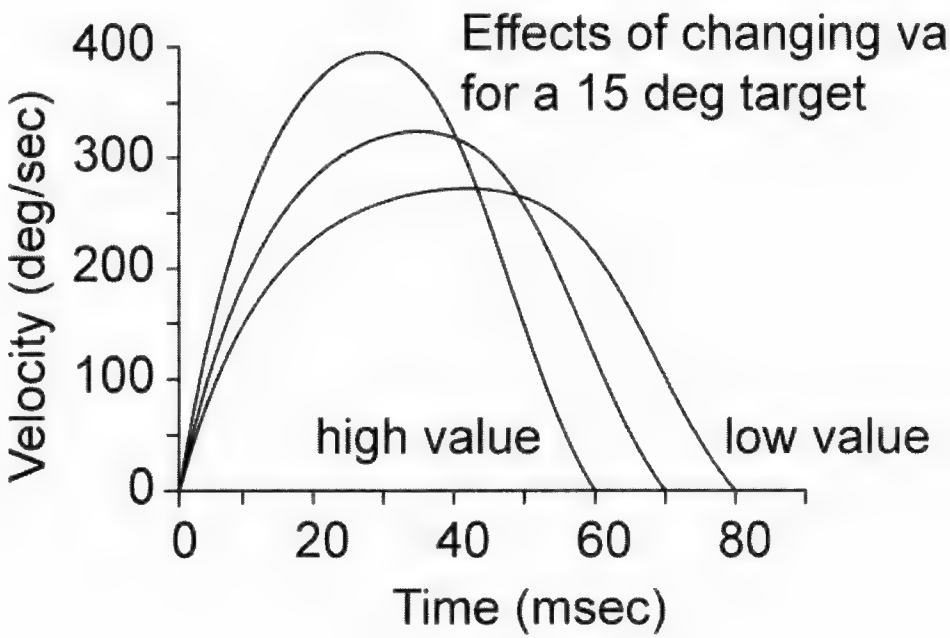


Prediction: If the value of saccade target changes, saccades must change their speeds and durations

Cost of a movement: $$J^{(0)} = \sum_{k=0}^{p-1} \mathbf{u}^{(k)T} L^{(k)} \mathbf{u}^{(k)} + \mathbf{y}^{(k+1)T} T^{(k+1)} \mathbf{y}^{(k+1)}$$

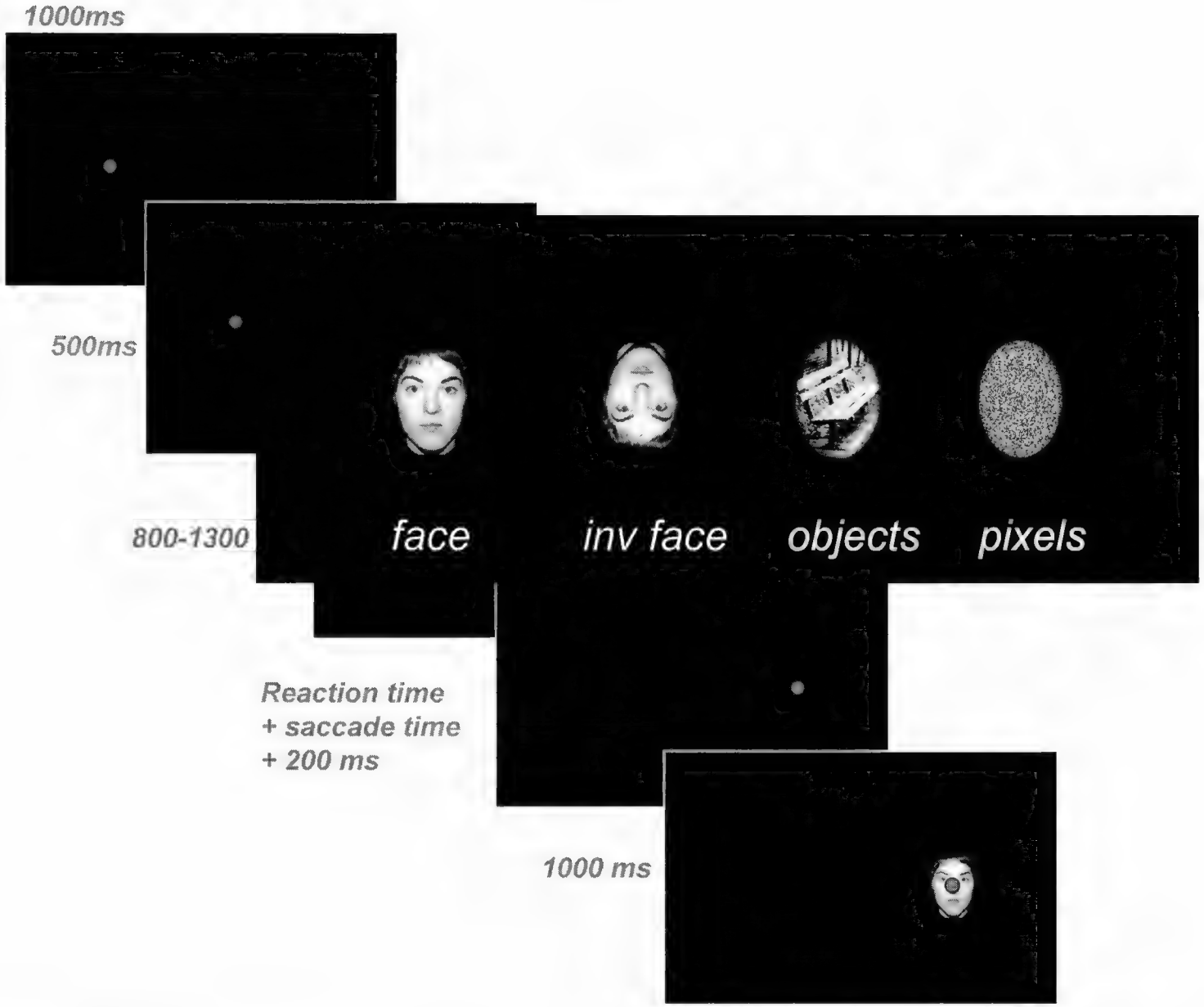
"cost of the motor commands"

"value of the target state"



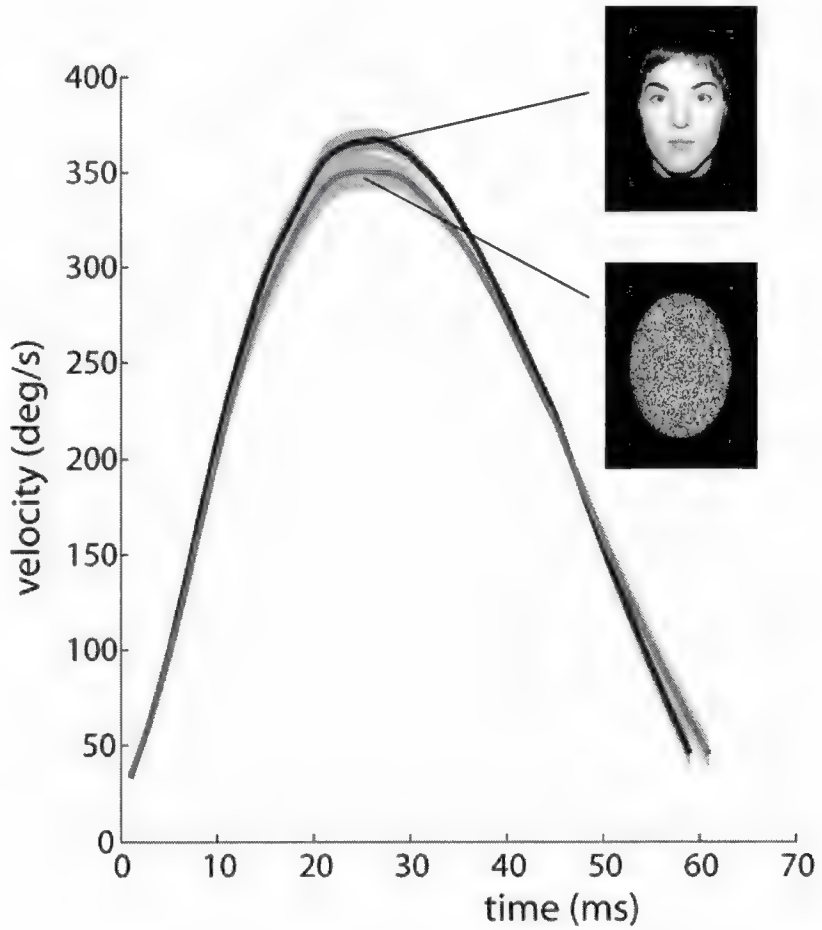
Chen-Harris et al. J Neurosci (2008)

Do saccade targets carry an intrinsic value?



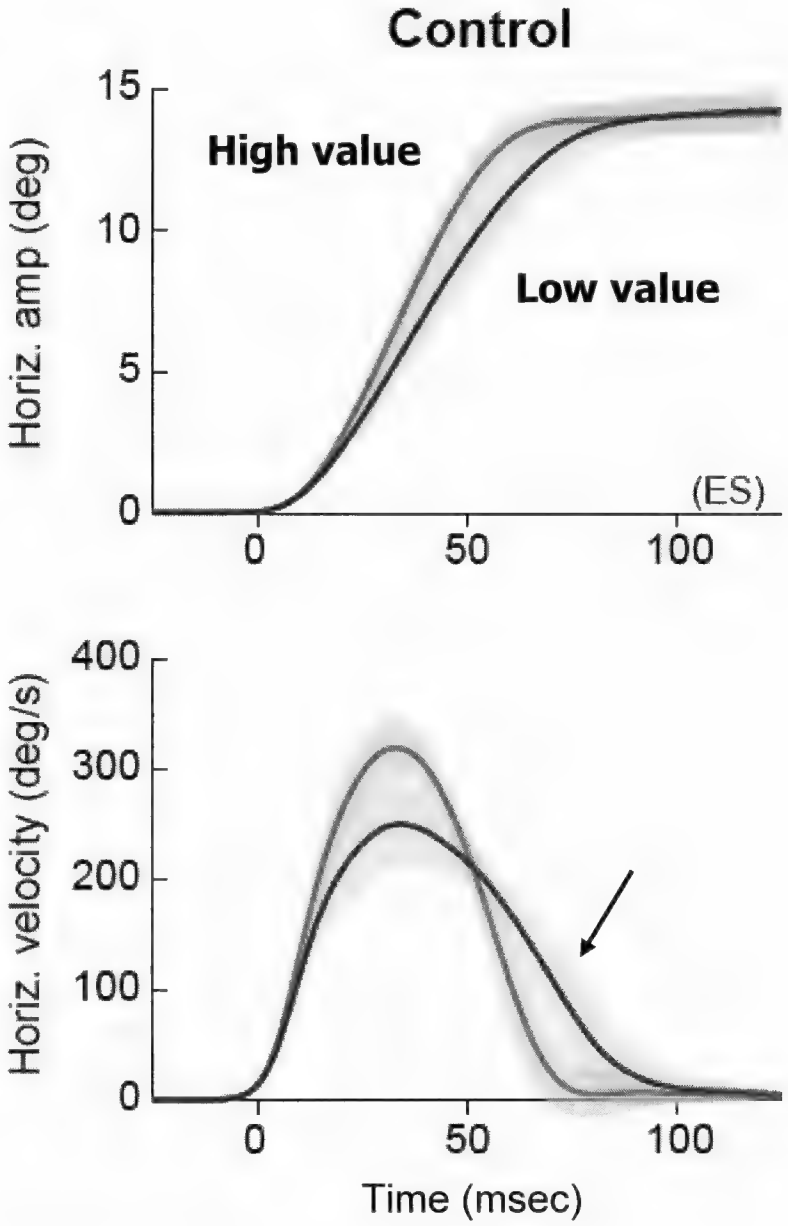
Minnan Xu-Wilson et al. (2009)

Saccades to faces are faster and shorter in duration



Minnan Xu-Wilson et al. (2009)

Saccade endpoints remain immune to changes in value because of the 'steering' via the cerebellum



Minnan Xu (unpublished data)

Summary

Minnan Xu

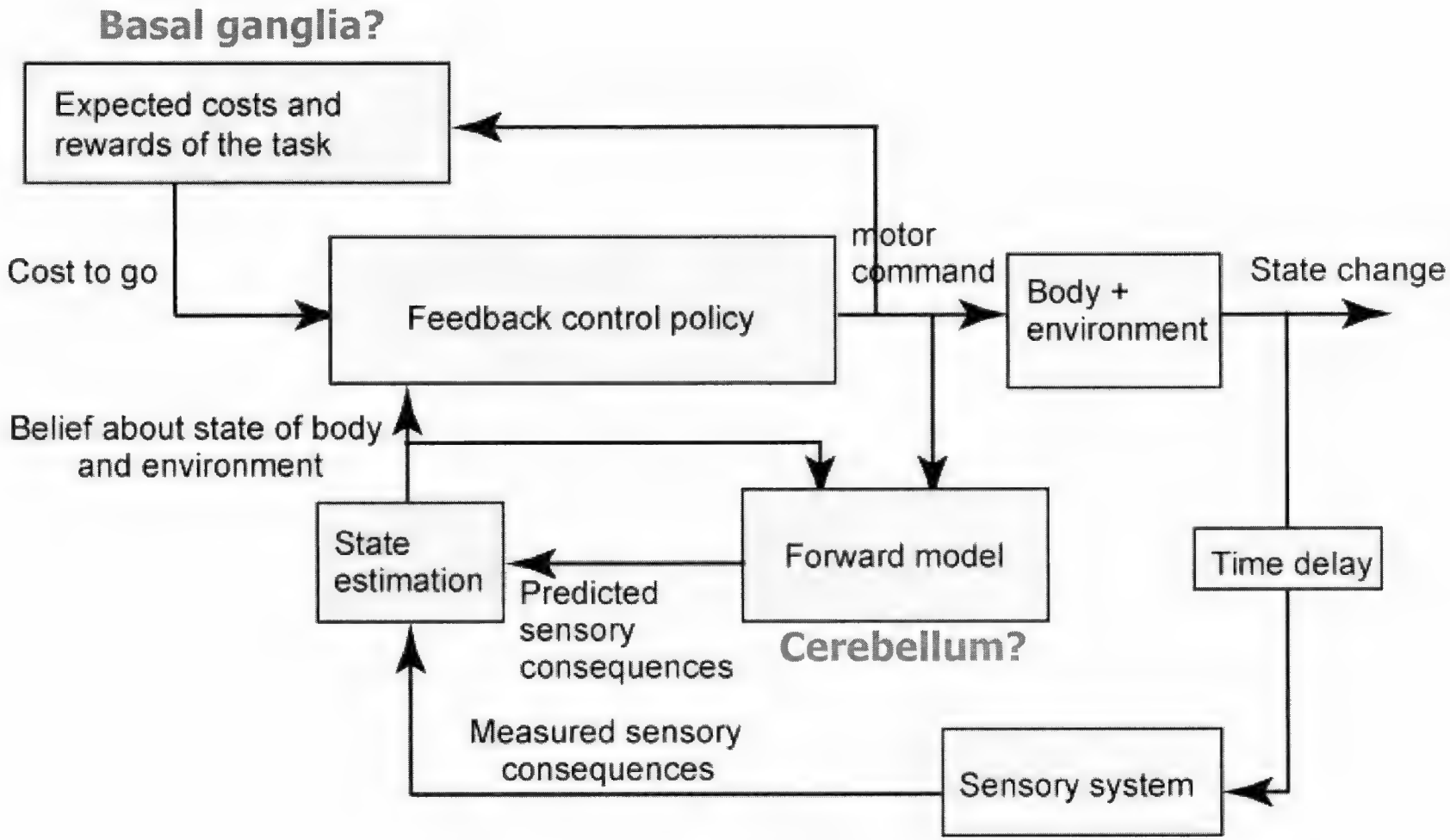
The 'value' of a stimulus affects trajectory of movements

Saccades remain accurate because of internal feedback that 'steers' the eyes to the target.

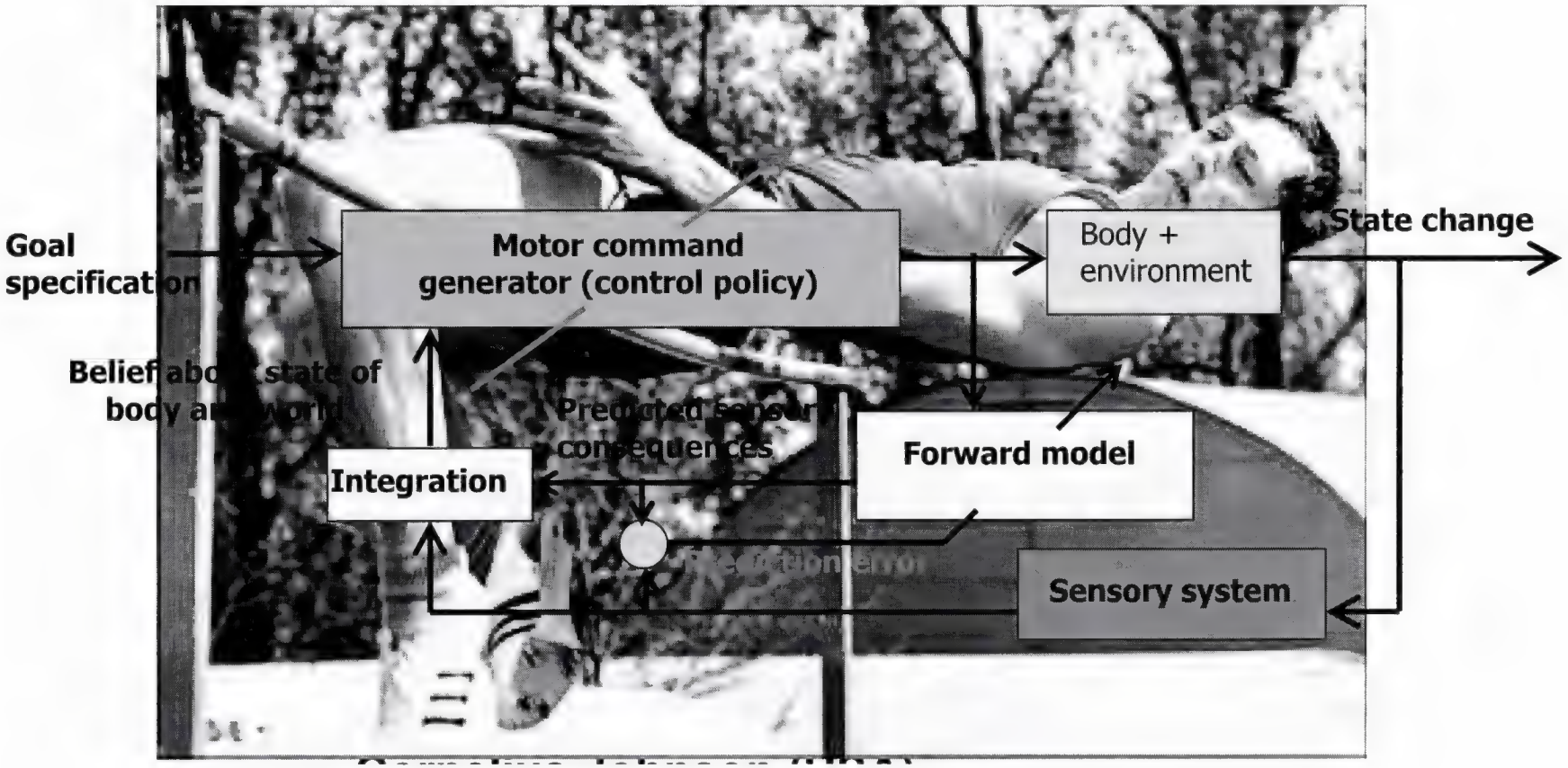
The internal feedback likely depends on the cerebellum.

What is motor control?

1. Learning to accurately predict the sensory consequences of our actions.
2. The search for a control policy that maximizes reward while expending as little effort as possible, i.e., a process of optimization.

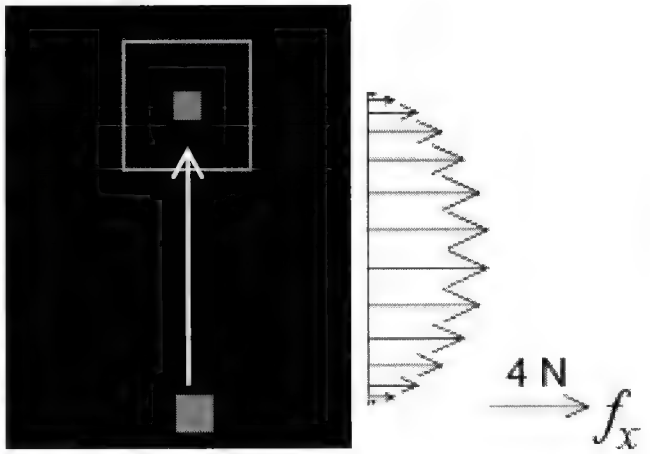


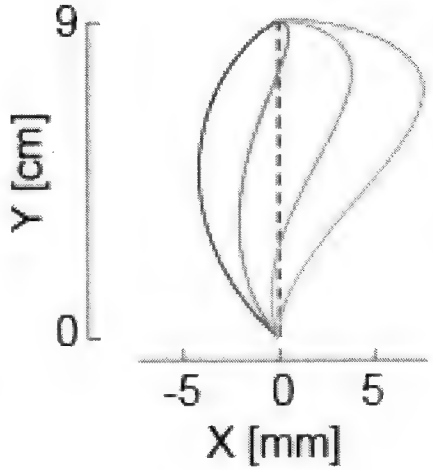
Motor learning as a process of optimization

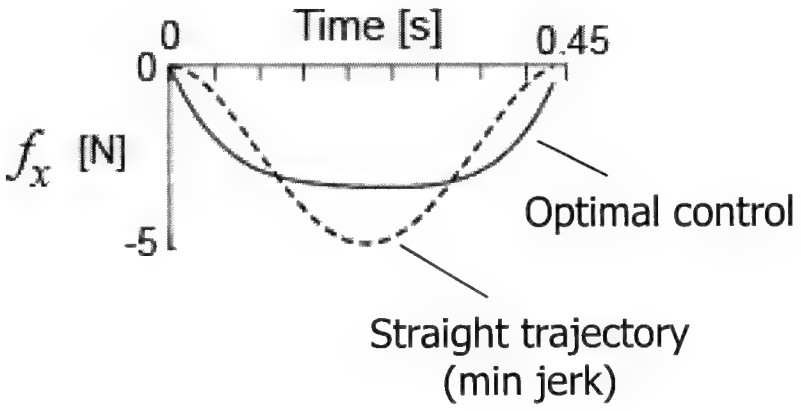


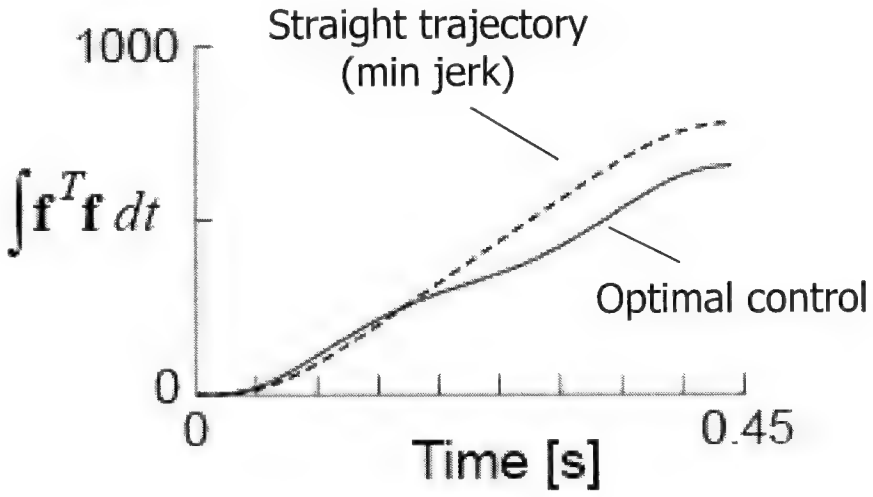
Dick Fosbury (USA)
Gold Medal, 1968 Olympics

Adaptation as a process of re-optimization



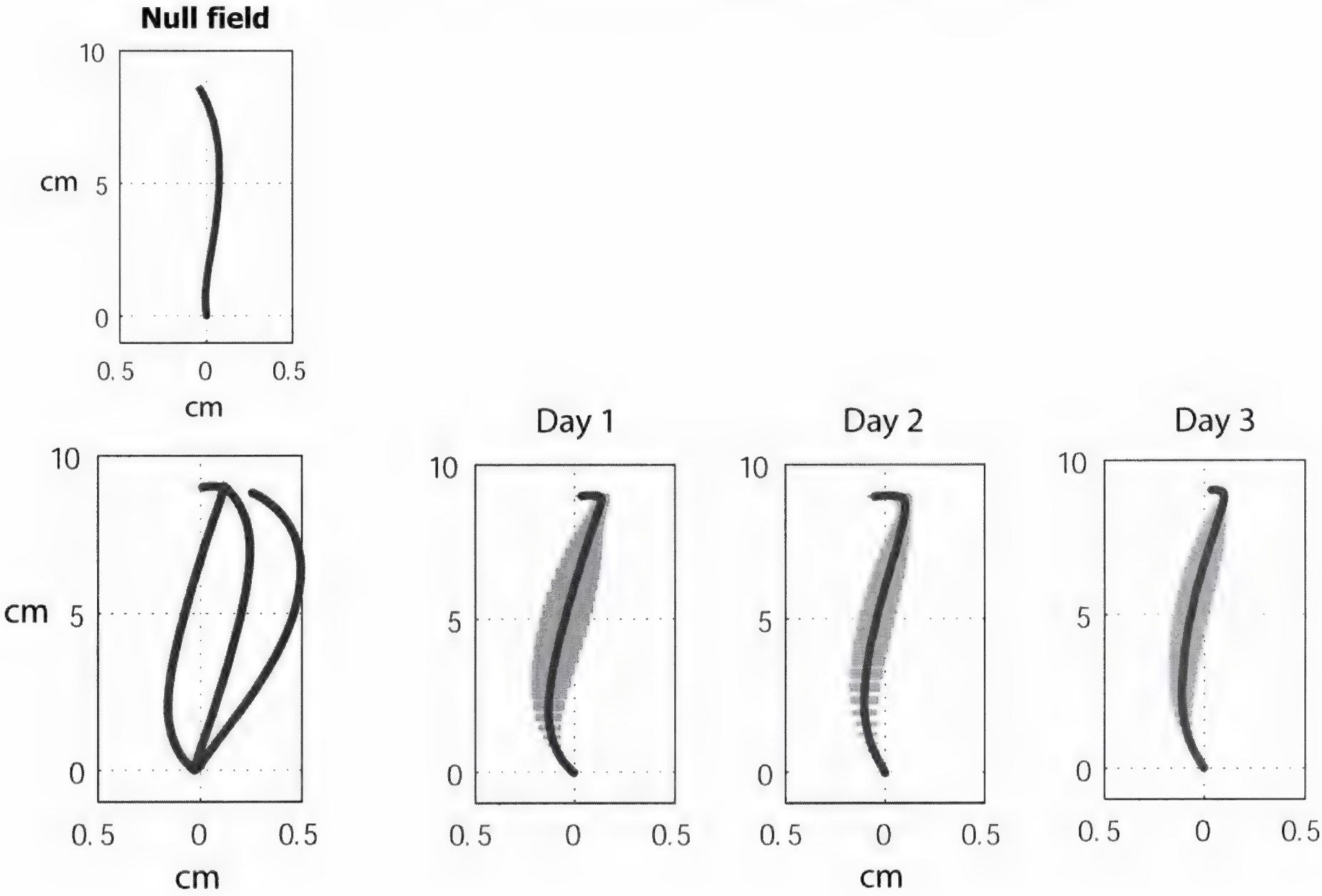






Izawa et al., J Neurosci 2008

Adaptation as a process of re-optimization



Izawa et al., J Neurosci 2008

Summary

Jun Izawa

Motor adaptation is a process of re-optimization:

Our brain learns to accurately predict the sensory consequences of motor commands (forward model).

Our brain uses that model to maximize rewards while minimizing effort.

Summary:

Motor control is about solving two distinct problems:

- What are the sensory consequences of my motor commands?
- What motor commands should I produce to maximize reward?

Adaptation experiments engages multiple learning systems:

- A fast learning system that has poor retention.
- A slow learning system that hardly forgets.

Cerebellar subjects show a particular deficit in the fast system.

The fast learning system appears to be in the forward model: the system that monitors motor commands and corrects them as they are generated.

Cerebellar subjects show a particular deficit in the forward model, as evidenced by their inability to compensate for variability caused by changes in the value of the stimulus.

Optimal control of movement

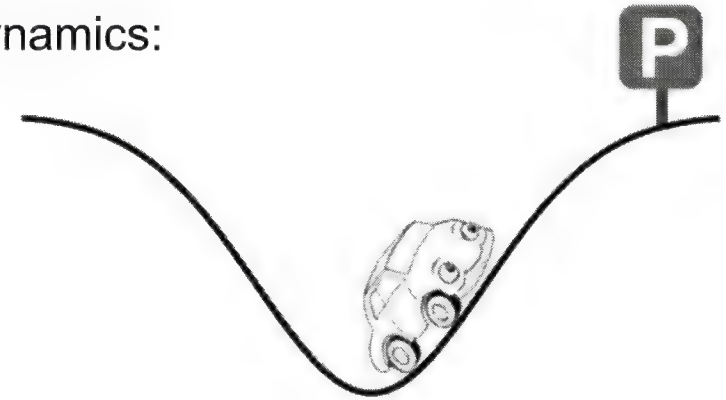
Emo Todorov

Departments of Applied Mathematics
and Computer Science & Engineering

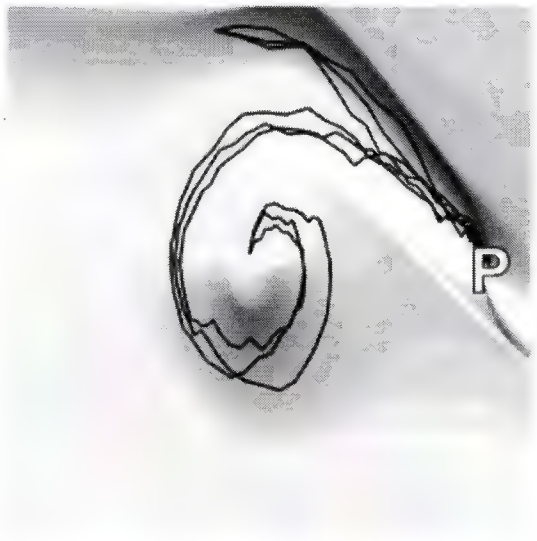
University of Washington

Simple assumptions yield rich predictions

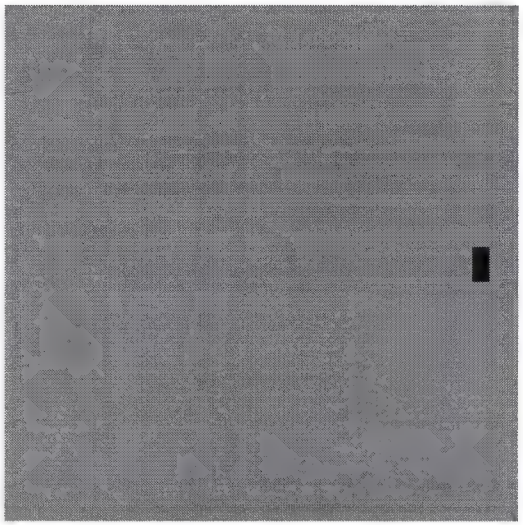
dynamics:



optimal control law:

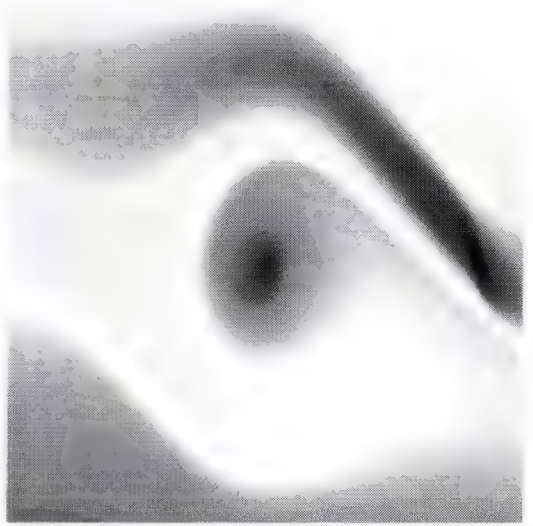


cost function:



+ control^2

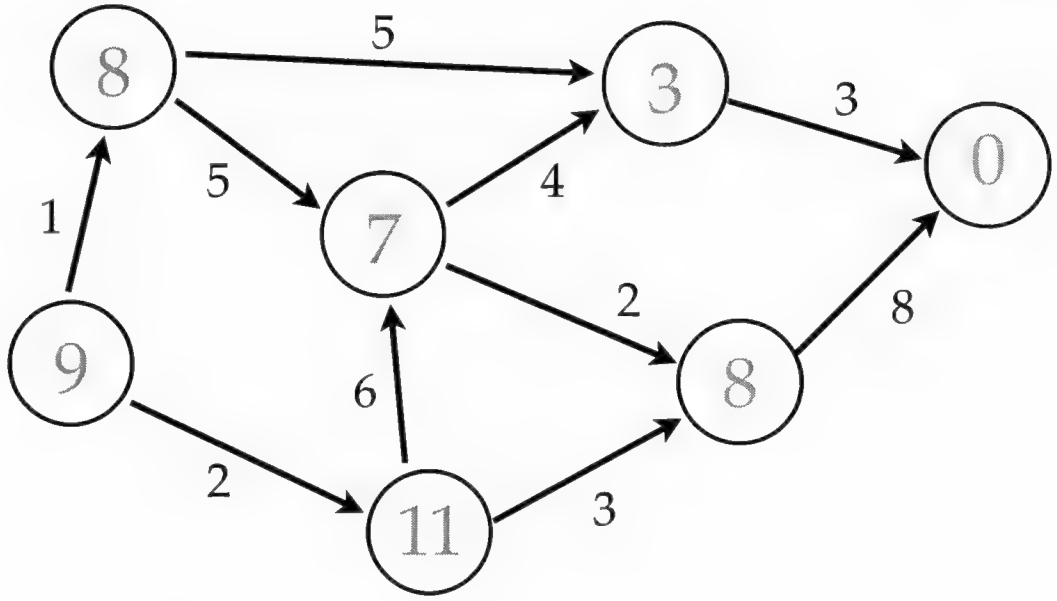
optimal cost-to-go function:



How to solve optimal control problems

dynamics:

$$x' = f(x, u)$$

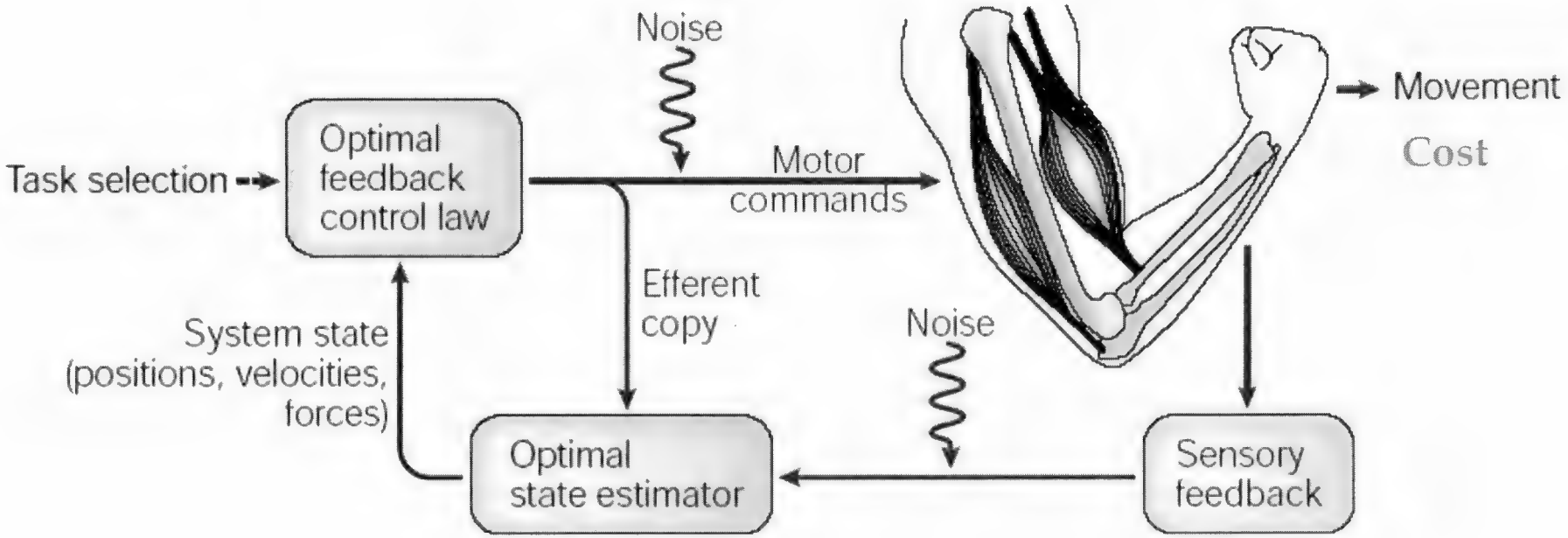


Bellman equation:

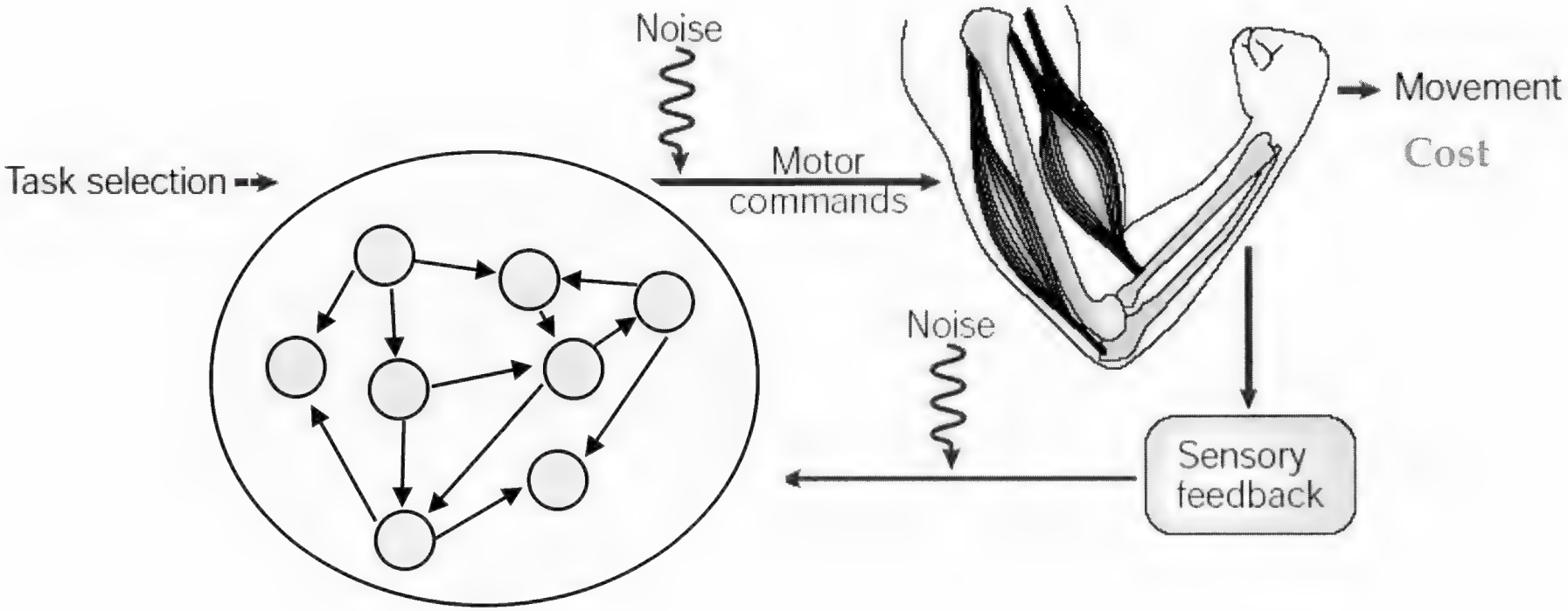
$$v(x) = \min_u \{l(x, u) + v(f(x, u))\}$$

(approximate) dynamic programming algorithms

Bayesian estimation and optimal control

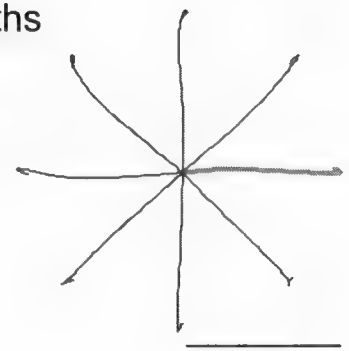


Bayesian estimation and optimal control

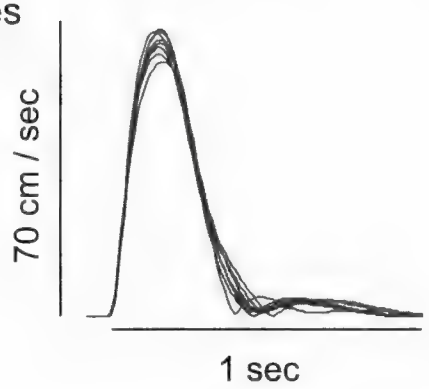


Neural networks for reaching

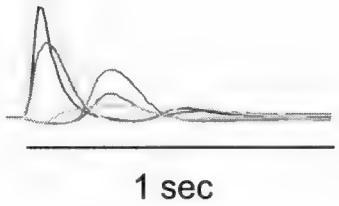
hand paths



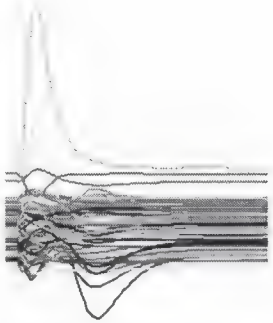
speed profiles



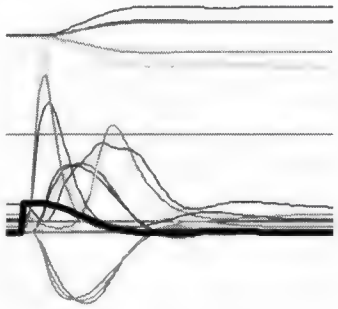
muscles



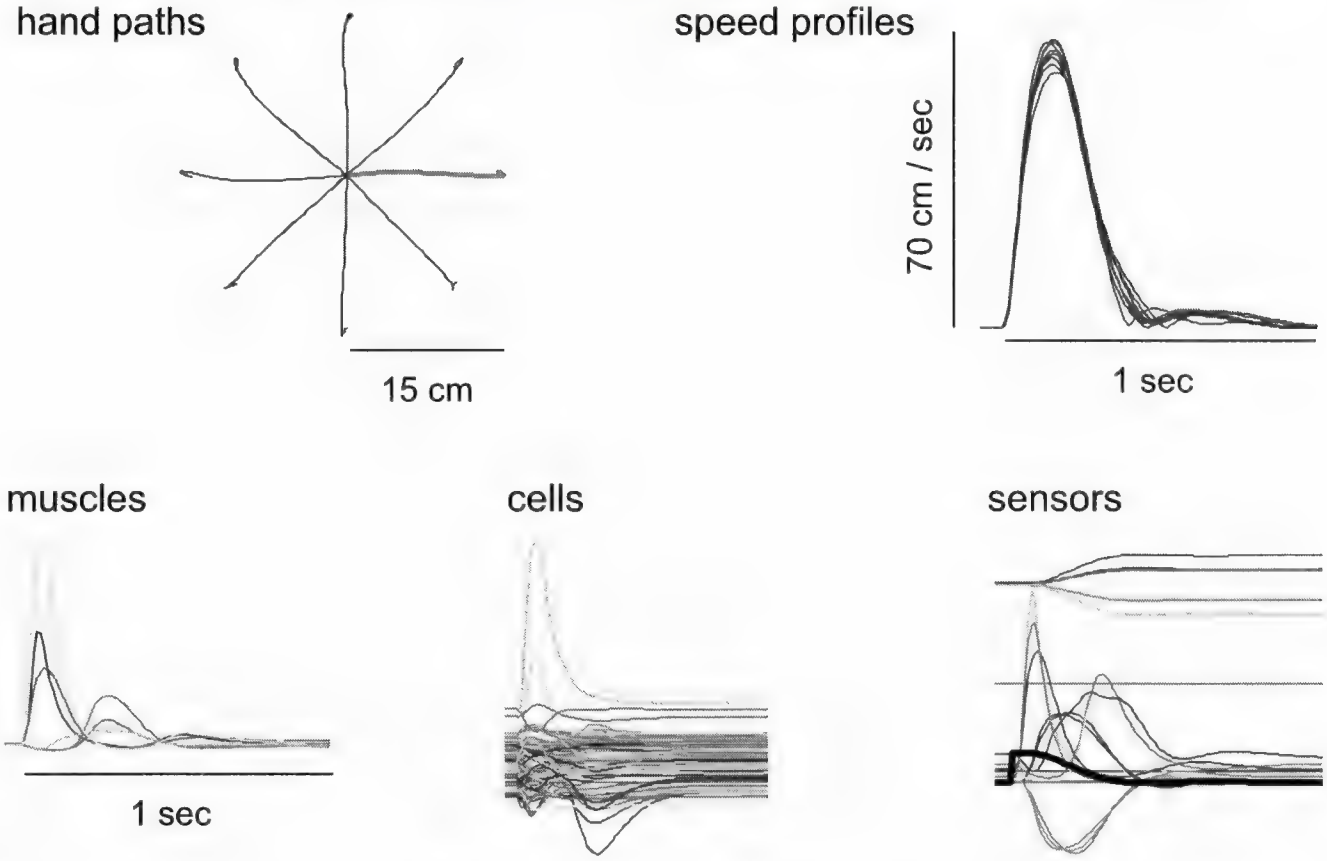
cells



sensors

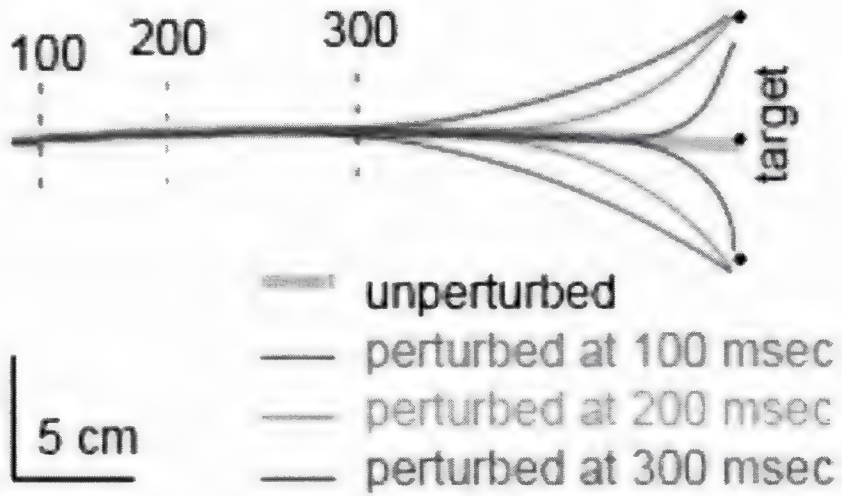


Neural networks for reaching

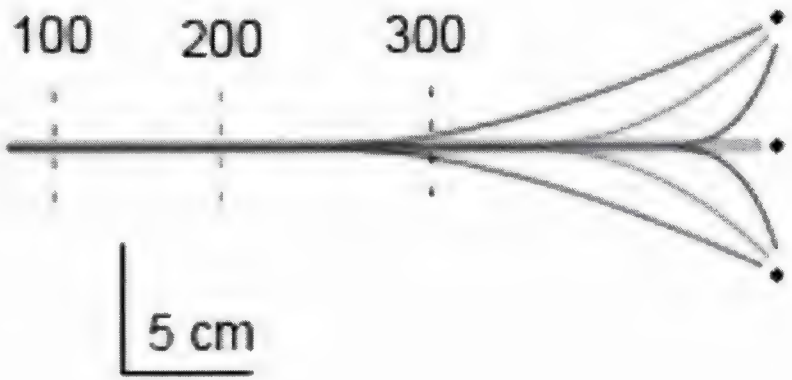


Incomplete error correction can be optimal

data

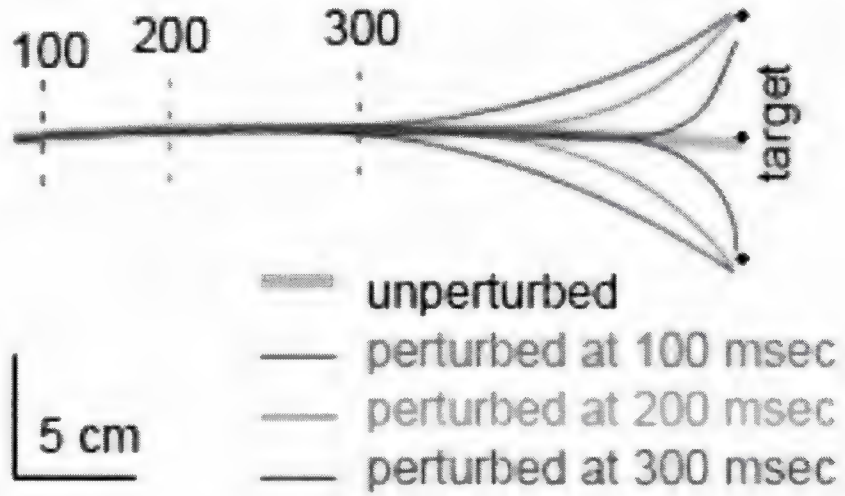


model

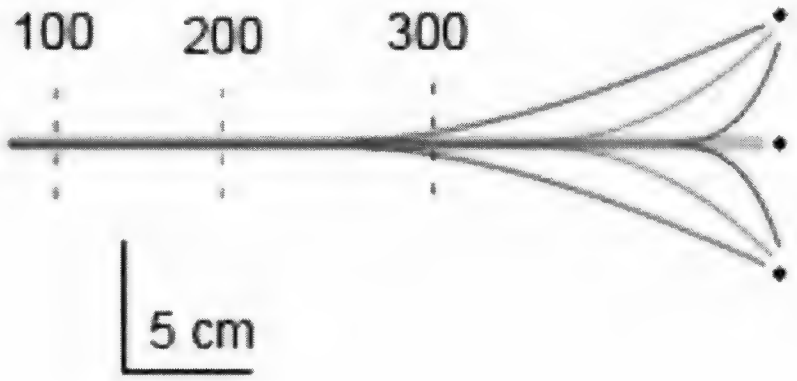


Incomplete error correction can be optimal

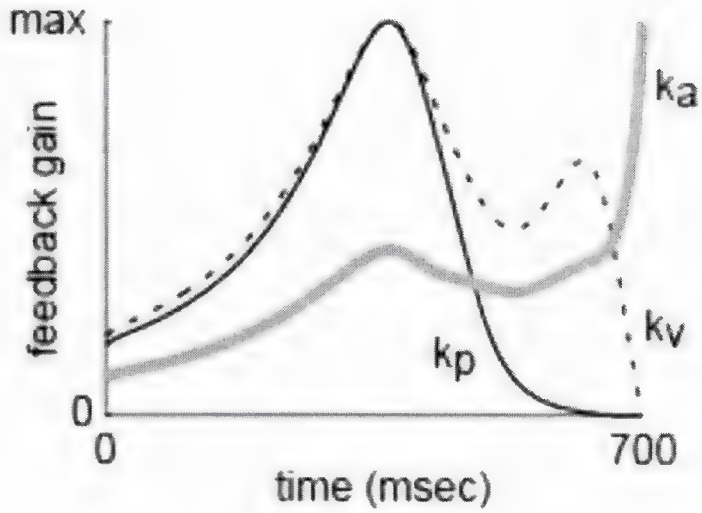
data



model

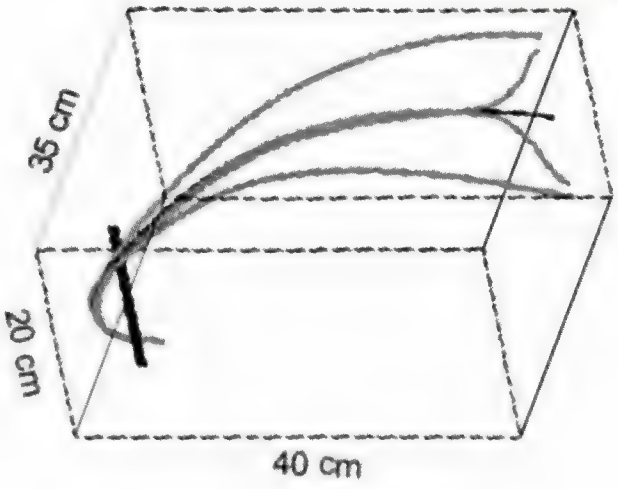


explanation:
stability-accuracy trade-off

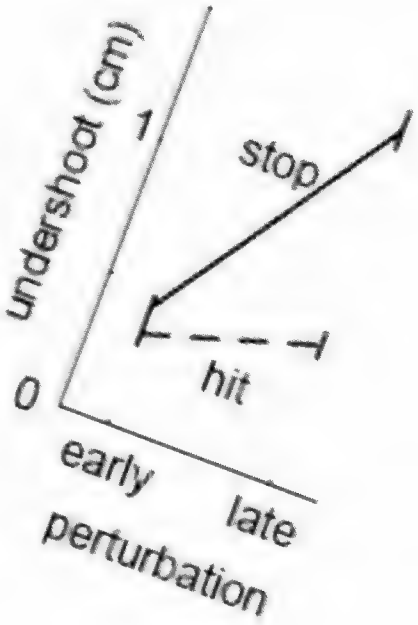


prediction:
undershoot will decrease if
stopping is not required

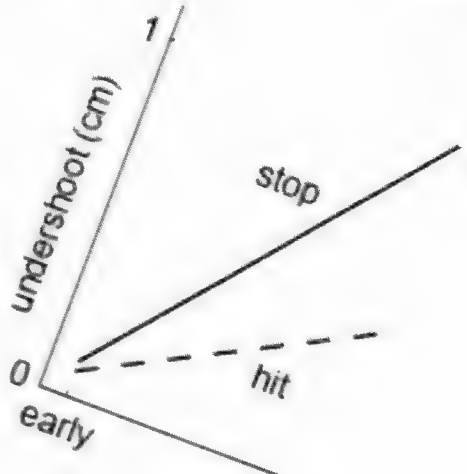
Testing the prediction



data



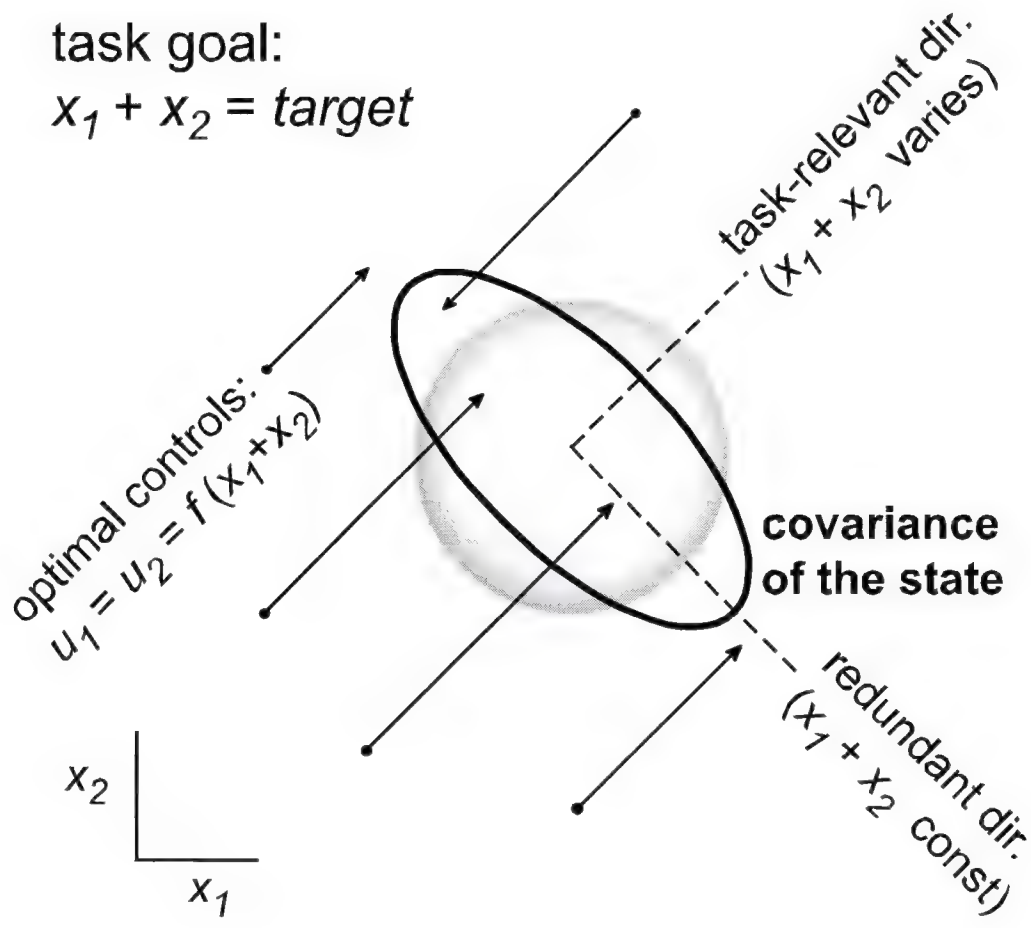
model



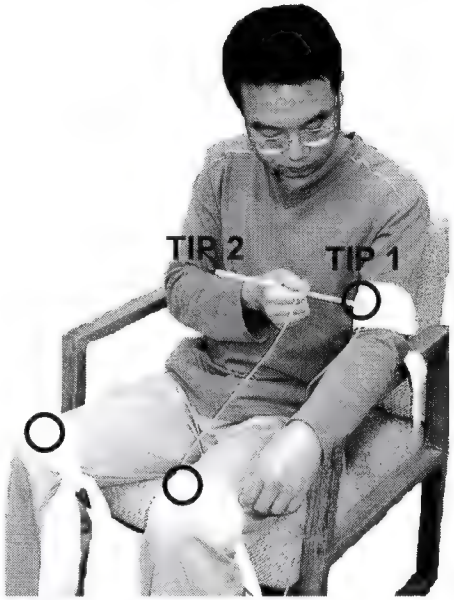
Minimal intervention principle

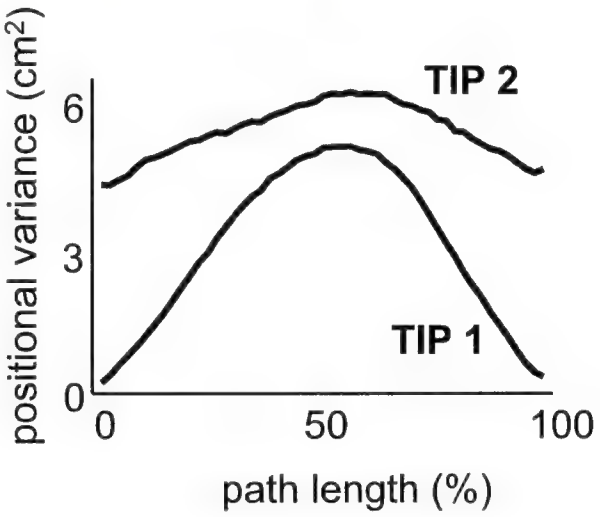
Optimal feedback controllers suppress task-relevant errors, while allowing fluctuations in redundant dimensions

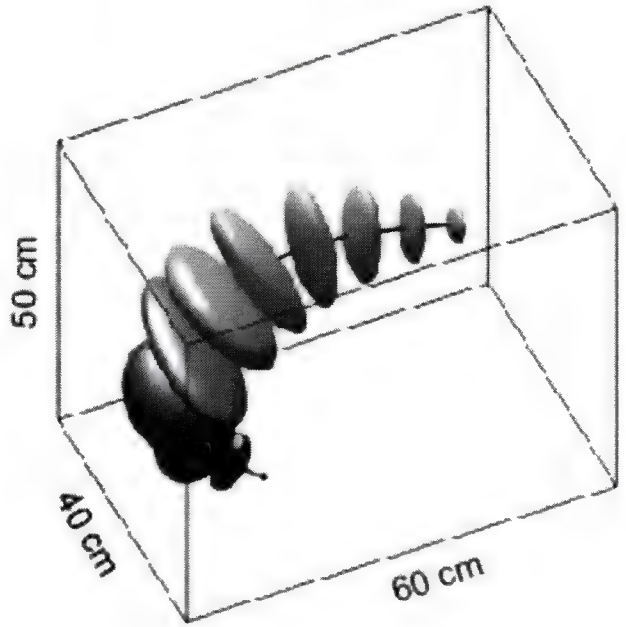
The optimal control signals are coupled - which implies a form of hierarchical control



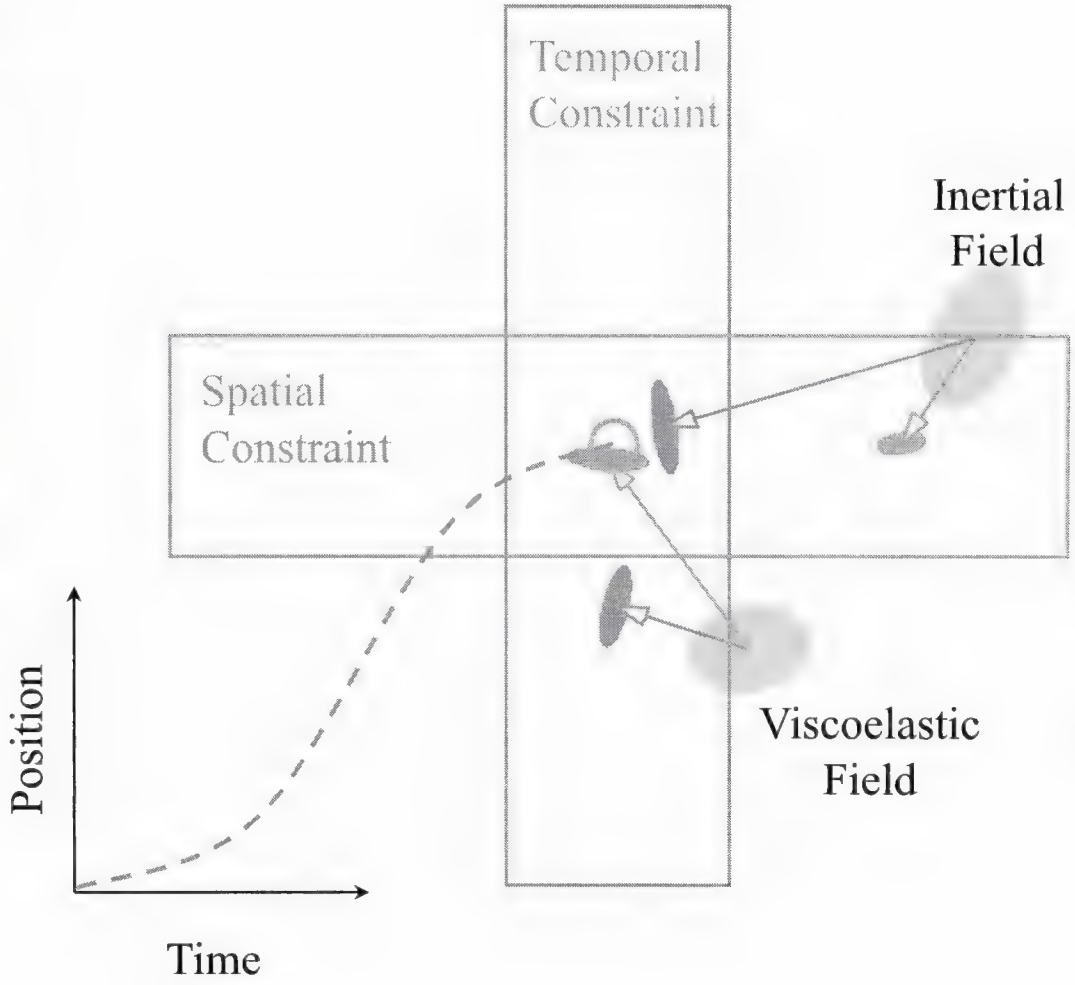
Increased variability in redundant dimensions



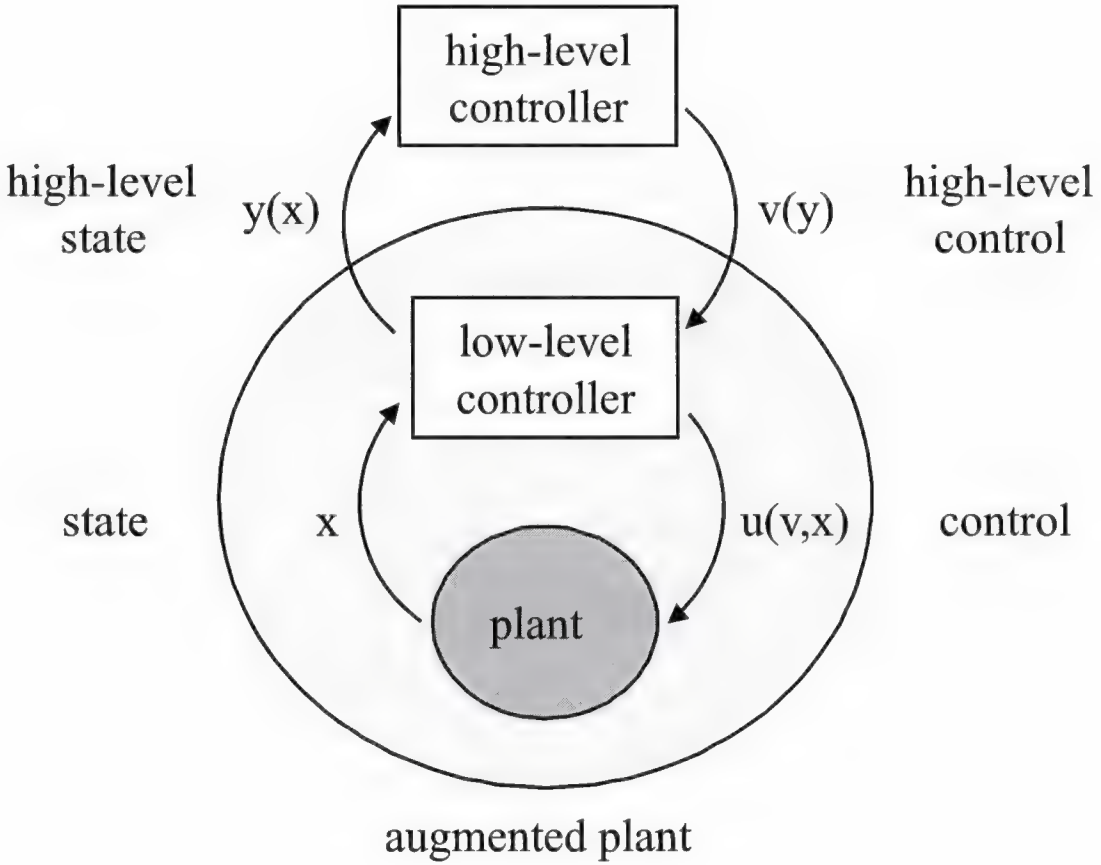




Reduced adaptation in redundant dimensions



Incorporating hierarchical structure a priori

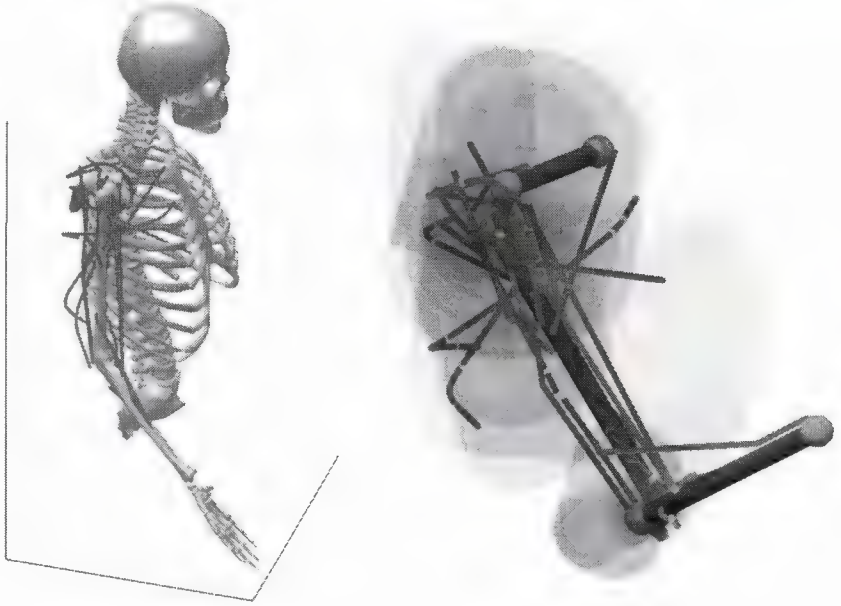


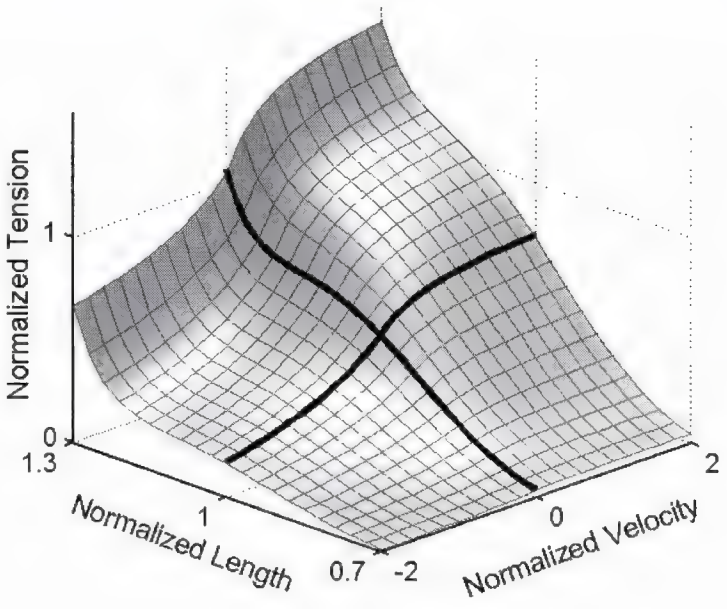
optimal feedback control in task space

feedback transformation making the plant "easier" to control

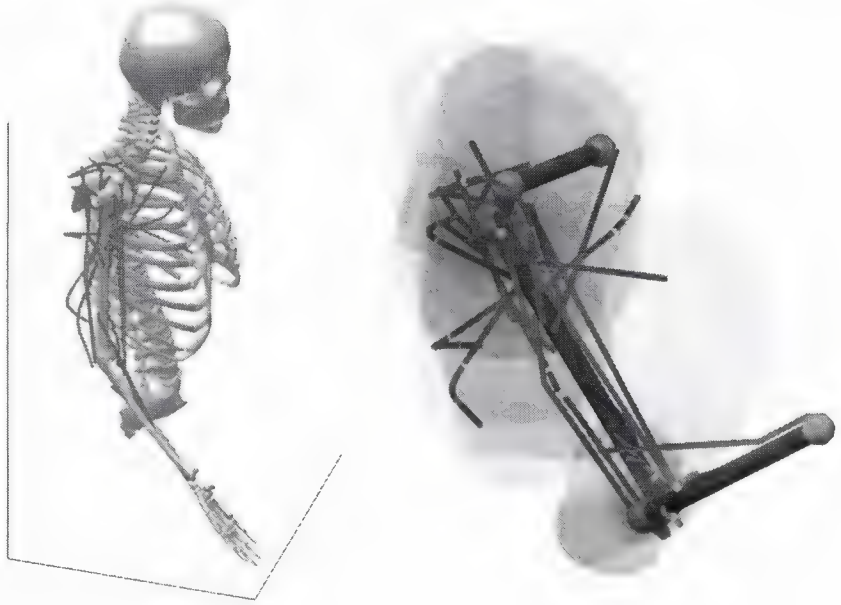
u(v,x) can be computed efficiently, by solving a linear equation

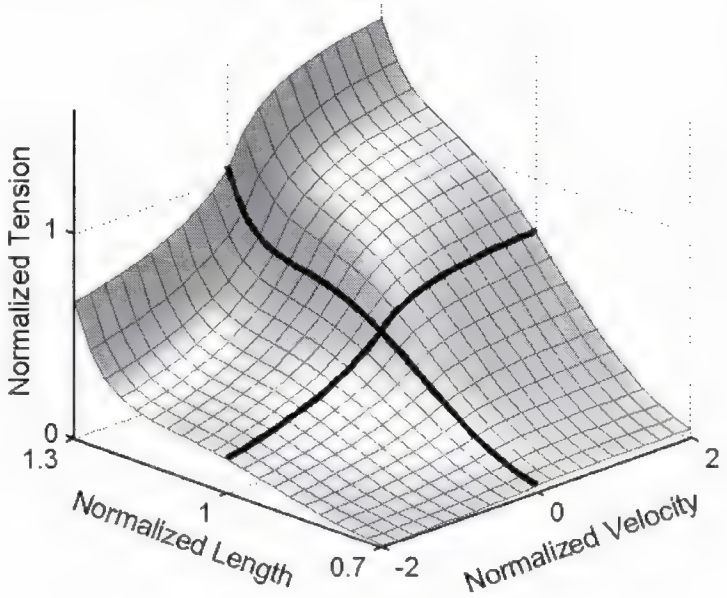
Application to complex biomechanical systems



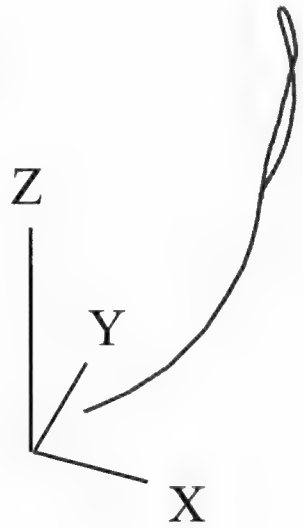


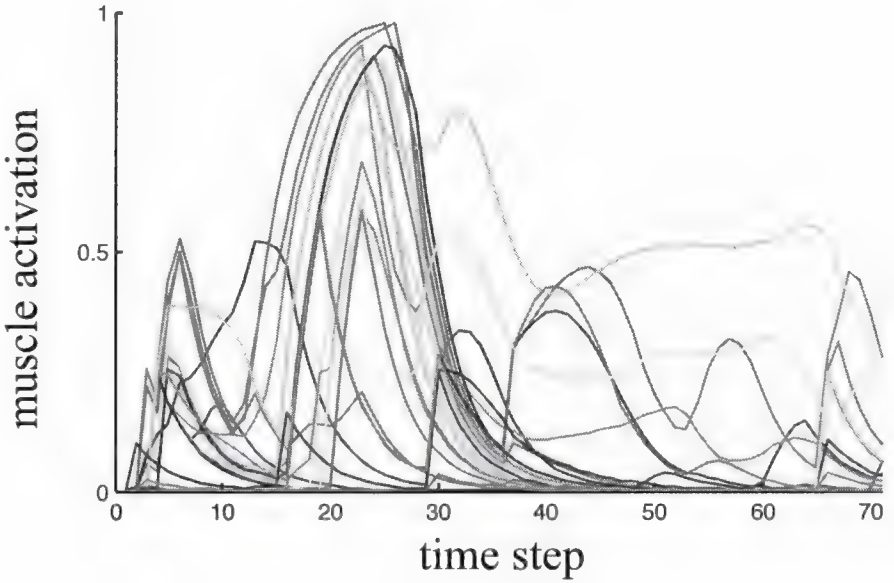
Application to complex biomechanical systems





hand path:





A more tractable formulation of optimal control

The passive dynamics in the absence of control are $p(x' \mid x)$

The controller can impose any dynamics $u(x' \mid x)$,
with the restriction that $u(x' \mid x) = 0$ whenever $p(x' \mid x) = 0$

The control cost is the KL divergence between u and p

The state cost (which defines the task goal) is $q(x)$

The passive dynamics p and state cost q can be arbitrary functions

Reducing optimal control to a linear problem

Desirability: $z(\mathbf{x}) = \exp(-v(\mathbf{x}))$ where $v(\mathbf{x})$ is the optimal cost-to-go

Next-state expectation: $\mathcal{G}[z](\mathbf{x}) = \int p(\mathbf{x}'|\mathbf{x})\, z(\mathbf{x}')\, d\mathbf{x}'$

Linear equation for z: $z(\mathbf{x}) = \exp(-q(\mathbf{x}))\, \mathcal{G}[z](\mathbf{x})$

Optimal control law: $u^*(\mathbf{x}'|\mathbf{x}) = \dfrac{p(\mathbf{x}'|\mathbf{x})\, z(\mathbf{x}')}{\mathcal{G}[z](\mathbf{x})}$

Reducing optimal control to a linear problem

Desirability: $z(\mathbf{x}) = \exp(-v(\mathbf{x}))$ where $v(\mathbf{x})$ is the optimal cost-to-go

Next-state expectation: $\mathcal{G}[z](\mathbf{x}) = \int p(\mathbf{x}'|\mathbf{x})\, z(\mathbf{x}')\, d\mathbf{x}'$

Linear equation for z: $z(\mathbf{x}) = \exp(-q(\mathbf{x}))\, \mathcal{G}[z](\mathbf{x})$

Optimal control law: $u^*(\mathbf{x}'|\mathbf{x}) = \dfrac{p(\mathbf{x}'|\mathbf{x})\, z(\mathbf{x}')}{\mathcal{G}[z](\mathbf{x})}$

Reducing optimal control to a linear problem

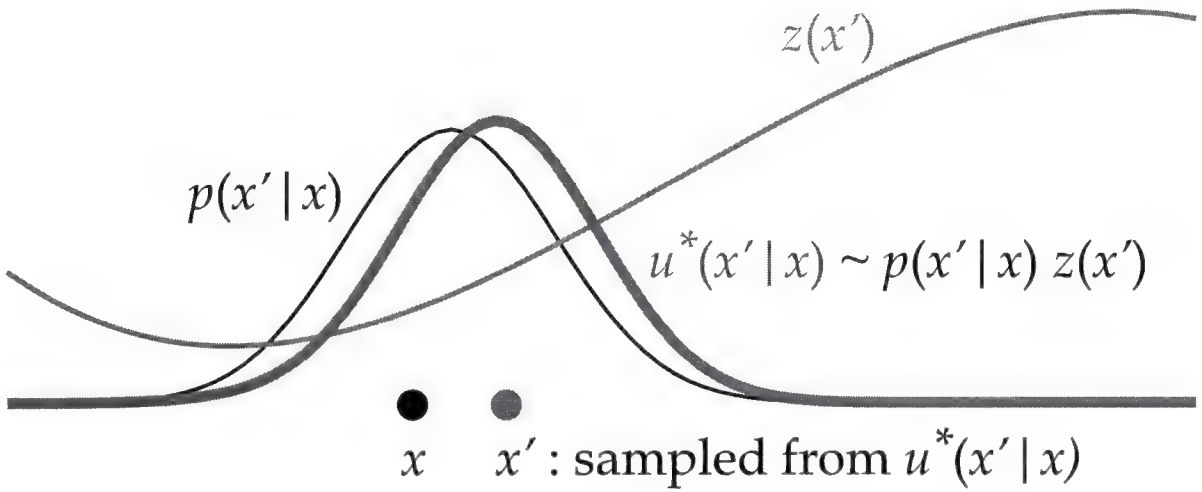
Desirability: $z(\mathbf{x}) = \exp(-v(\mathbf{x}))$ where $v(\mathbf{x})$ is the optimal cost-to-go

Next-state expectation: $\mathcal{G}[z](\mathbf{x}) = \int p(\mathbf{x}'|\mathbf{x})\, z(\mathbf{x}')\, d\mathbf{x}'$

Linear equation for z: $z(\mathbf{x}) = \exp(-q(\mathbf{x}))\, \mathcal{G}[z](\mathbf{x})$

Optimal control law: $u^*(\mathbf{x}'|\mathbf{x}) = \dfrac{p(\mathbf{x}'|\mathbf{x})\, z(\mathbf{x}')}{\mathcal{G}[z](\mathbf{x})}$

Multiplication by z pushes the system towards more desirable states:



Linearity implies compositionality

Consider a set of first-exit or finite-horizon problems which only differ by their final costs $g_i(\mathbf{x})$, and let the corresponding desirabilities be $z_i(\mathbf{x})$

If the final cost for a new (composite) problem can be expressed as

$$g(\mathbf{x}) = -\log \sum_i w_i \exp(-g_i(\mathbf{x}))$$

then the desirability for the new problem is a fixed linear mixture:

$$z(\mathbf{x}) = \sum_i w_i \, z_i(\mathbf{x})$$

The optimal control law becomes a state-dependent linear mixture:

<u>discrete time</u>

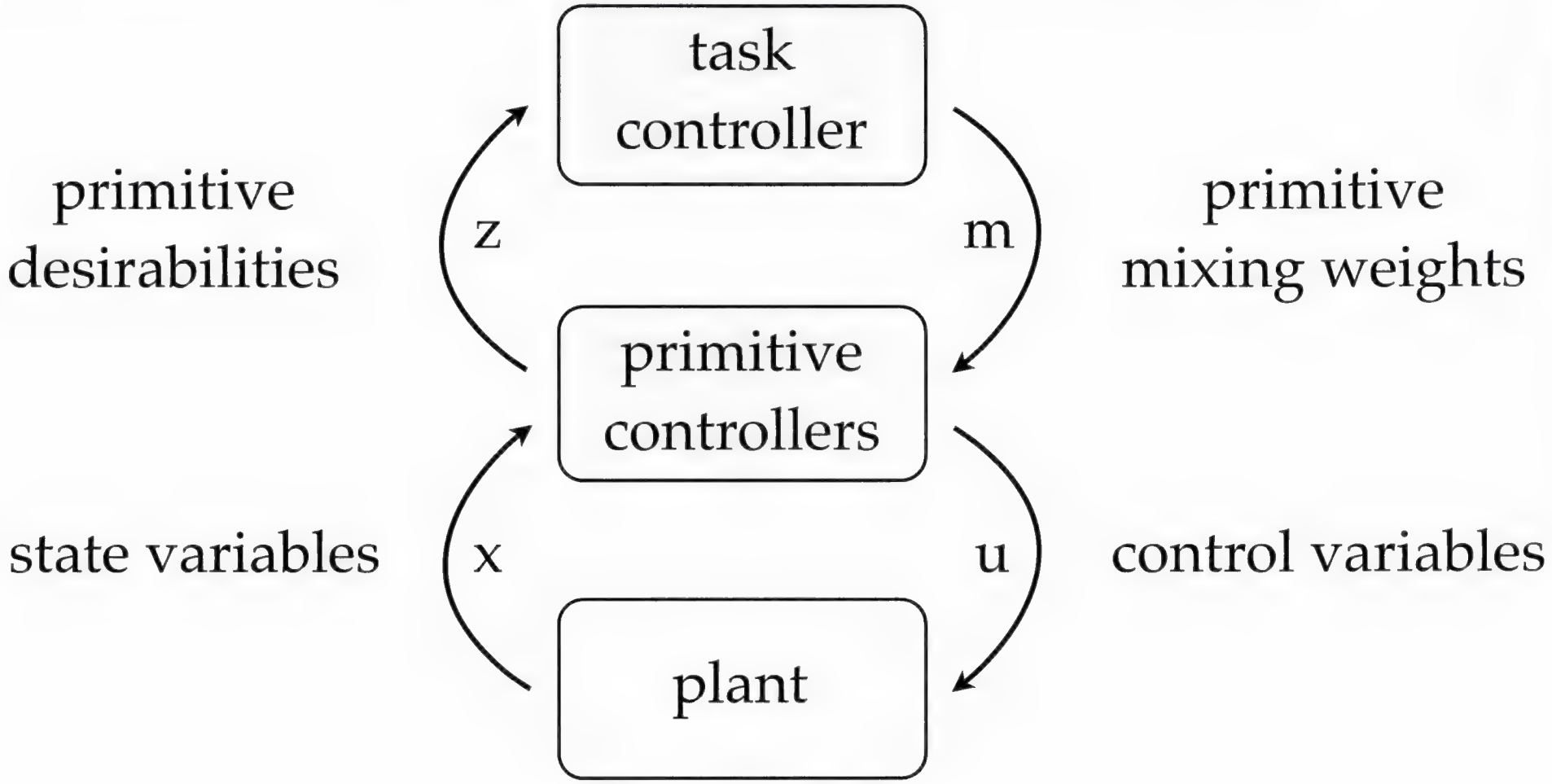
$$u^*(\mathbf{x}'|\mathbf{x}) = \sum_i m_i(\mathbf{x}) \, u_i^*(\mathbf{x}'|\mathbf{x})$$

<u>continuous time</u>

$$\mathbf{u}^*(\mathbf{x}) = \sum_i m_i(\mathbf{x}) \, \mathbf{u}_i^*(\mathbf{x})$$

$$m_i(\mathbf{x}) = \frac{w_i \, z_i(\mathbf{x})}{\sum_k w_k \, z_k(\mathbf{x})}$$

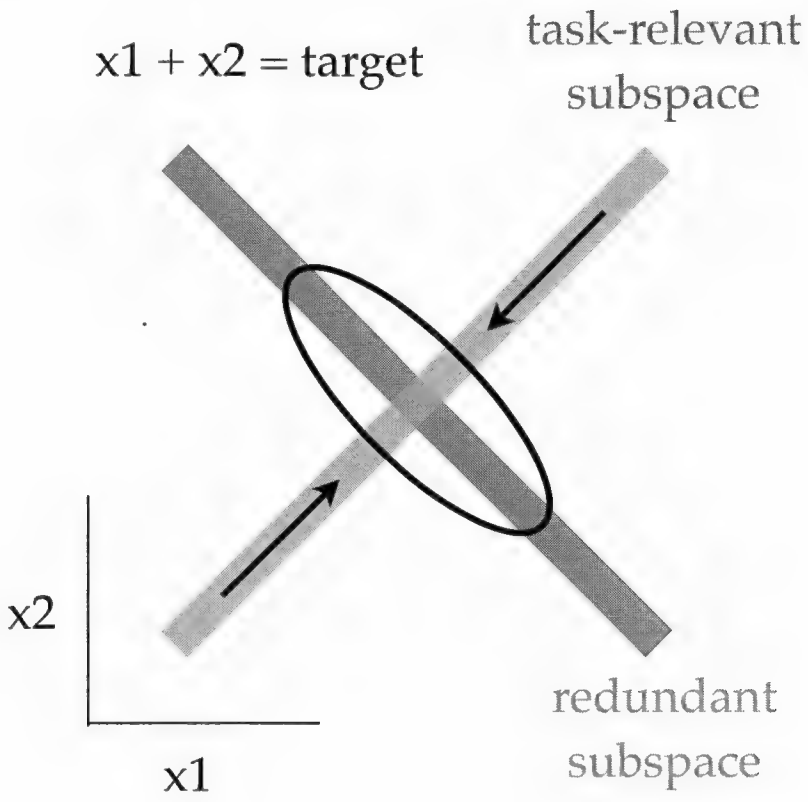
Another scheme for hierarchical control



the mixing weights are large for primitives whose desirability is high AND whose final cost is similar to the task cost (i.e. w is large)

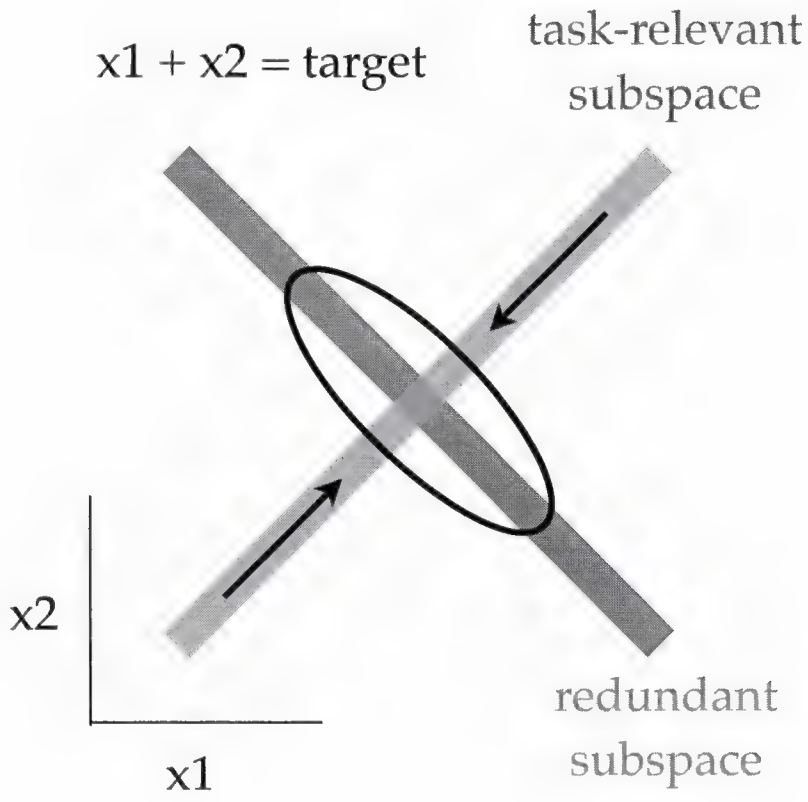
Dimensionality reduction or expansion? Both!

Optimal feedback control for redundant systems predicts lack of corrections in the redundant subspace

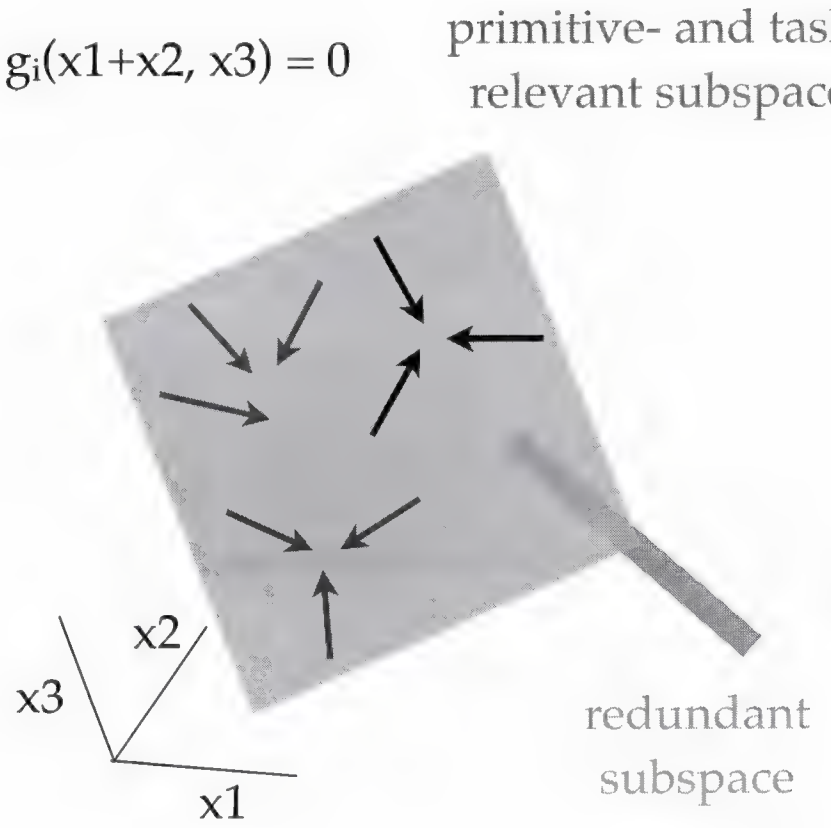


Dimensionality reduction or expansion? Both!

Optimal feedback control for redundant systems predicts lack of corrections in the redundant subspace



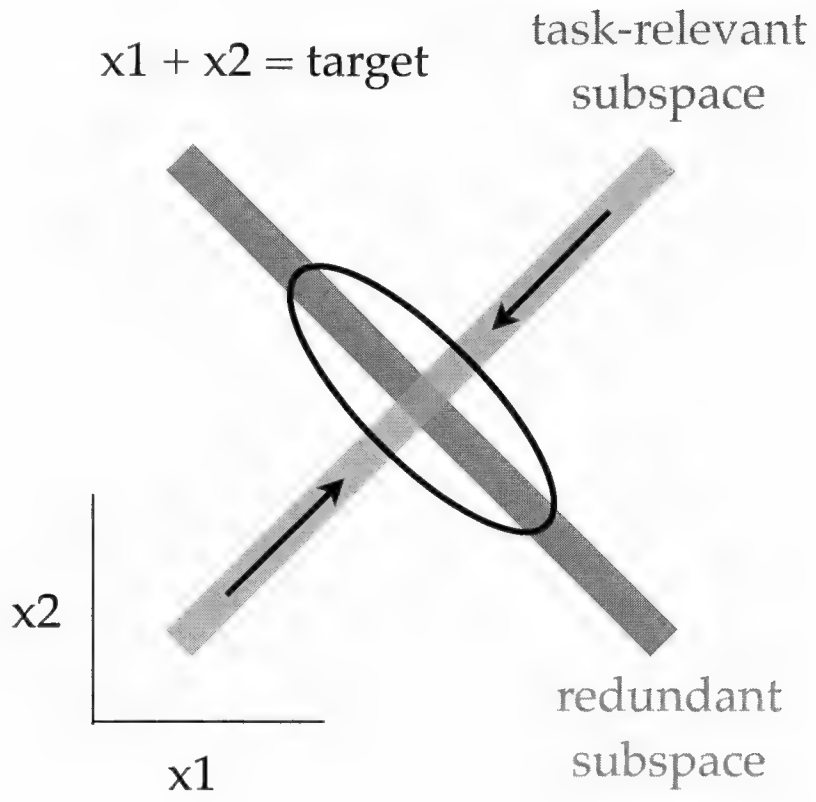
This is still true, however the primitives span the **space of feedback control laws** acting within the task-relevant subspace

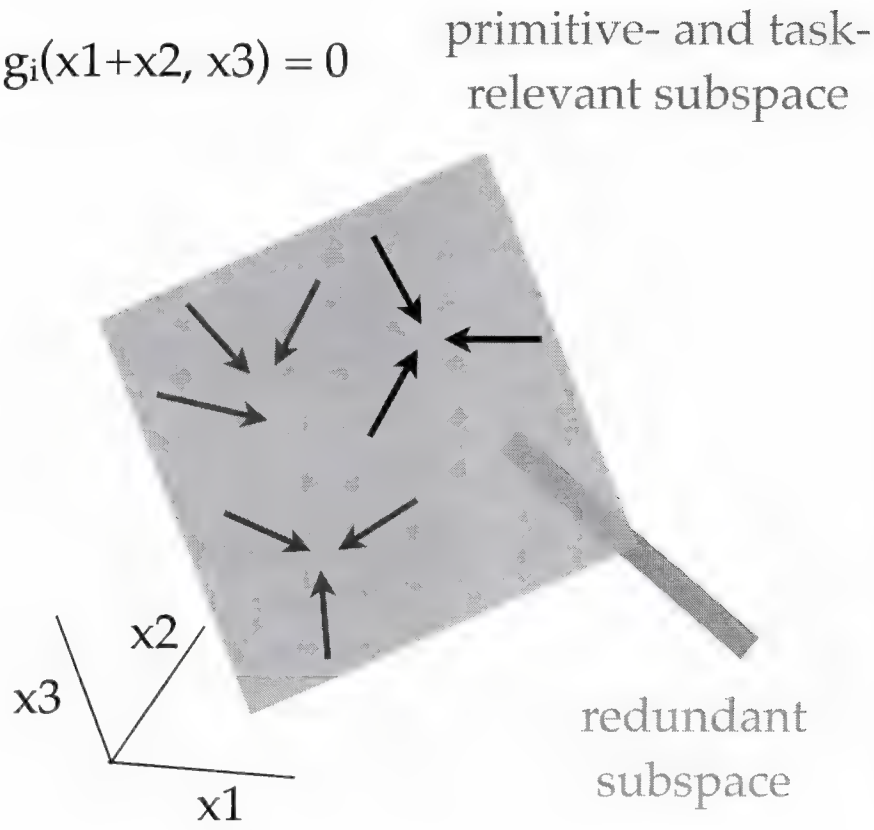


Dimensionality reduction or expansion? Both!

Optimal feedback control for redundant systems predicts lack of corrections in the redundant subspace

This is still true, however the primitives span the **space of feedback control laws** acting within the task-relevant subspace





Redundancy reduces the number of necessary primitives, but spanning even a subset of all feedback control laws requires more primitives than the number of state or control variables

Minimal basis sets of primitives

The mapping from the boundary to the interior is linear (Green's function)

We can make this mapping explicit by discretizing the state space:

$$\mathbf{z}_{\mathcal{I}} = \left(Q_{\mathcal{I}\mathcal{I}} - P_{\mathcal{I}\mathcal{I}}\right)^{-1} P_{\mathcal{I}\mathcal{B}}\, \mathbf{z}_{\mathcal{B}}$$

Minimal basis sets of primitives

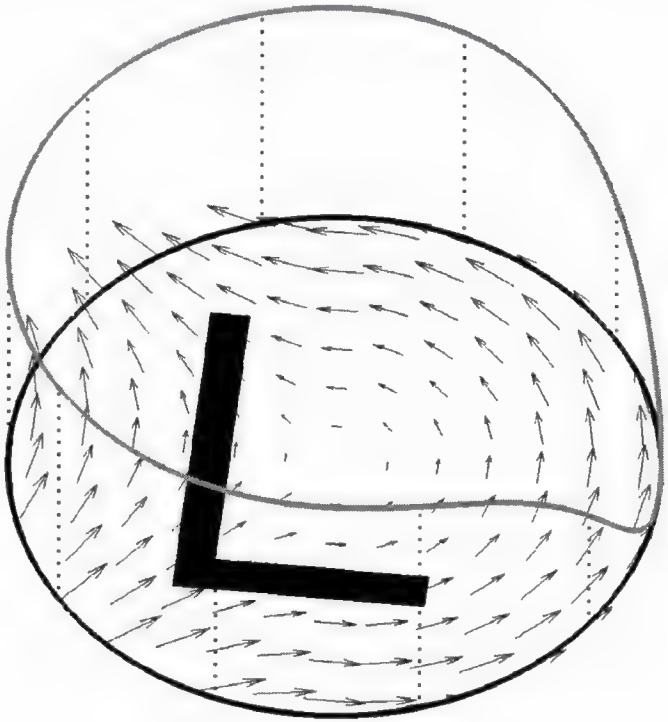
The mapping from the boundary to the interior is linear (Green's function)

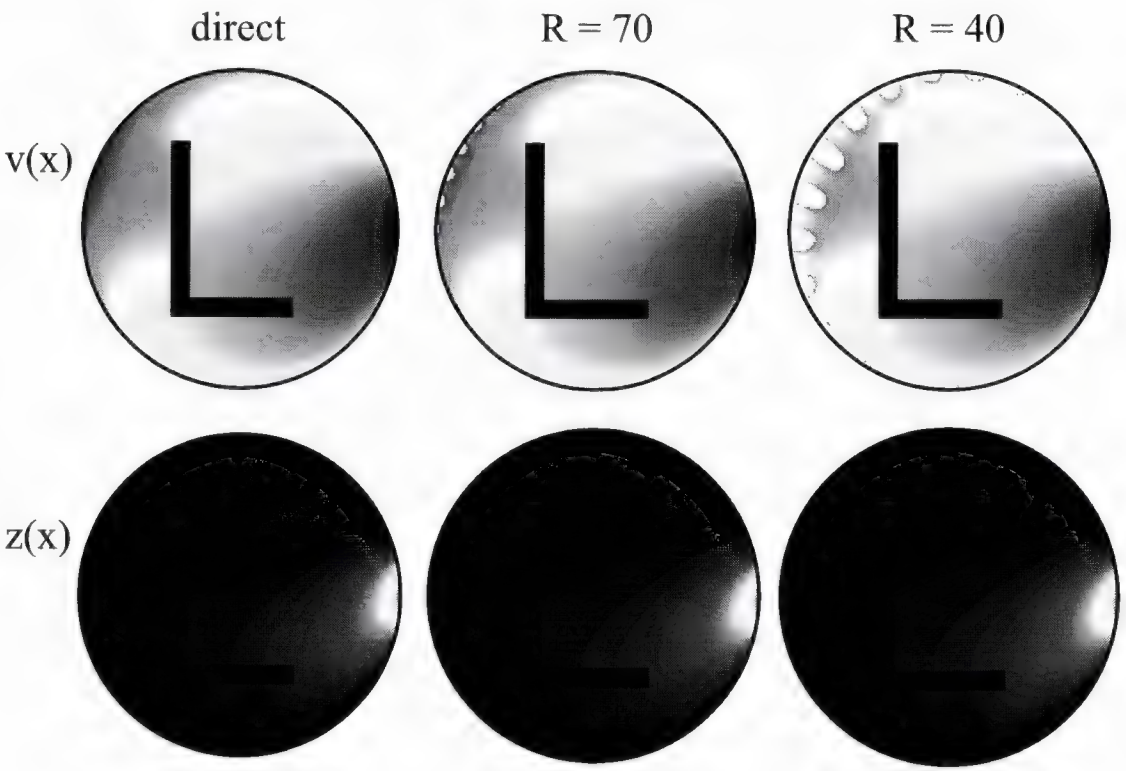
We can make this mapping explicit by discretizing the state space:

$$\mathbf{z}_{\mathcal{I}} = \left(Q_{\mathcal{I}\mathcal{I}} - P_{\mathcal{I}\mathcal{I}}\right)^{-1} P_{\mathcal{I}\mathcal{B}}\, \mathbf{z}_{\mathcal{B}}$$

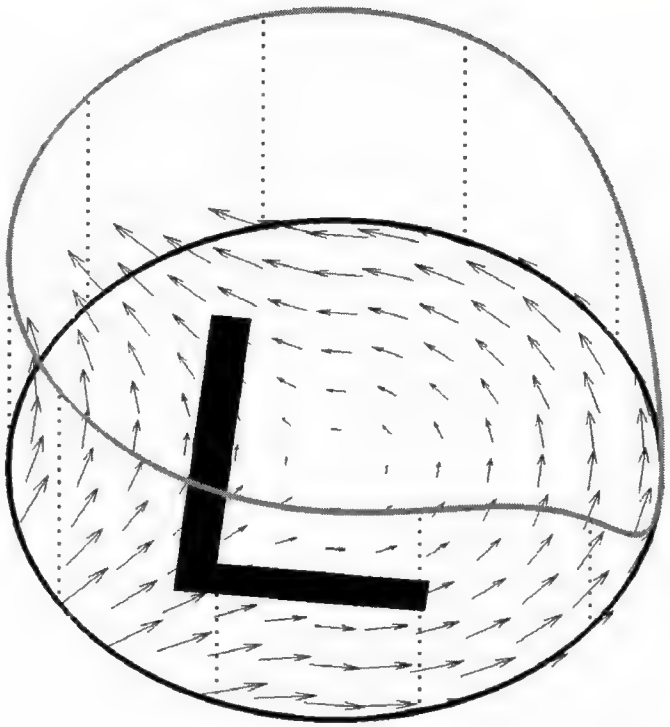
The singular vectors of this matrix provide the minimal basis set in terms of least-squares reconstruction of the desirability function

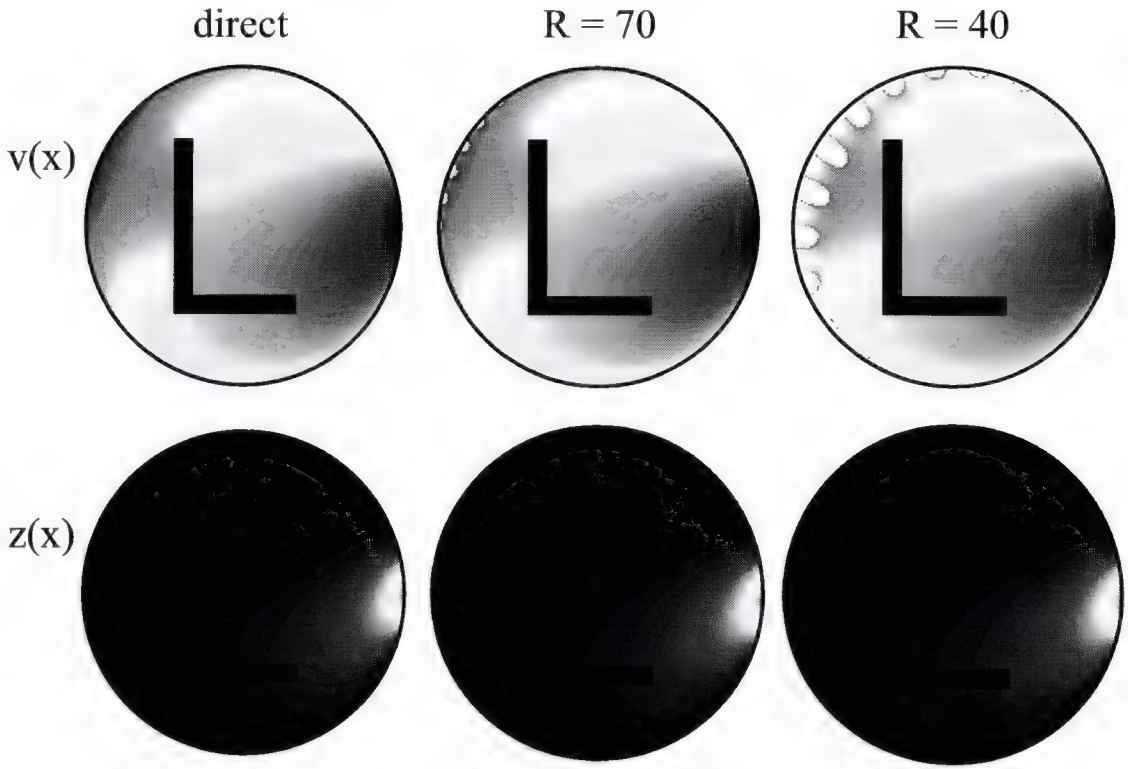
Example



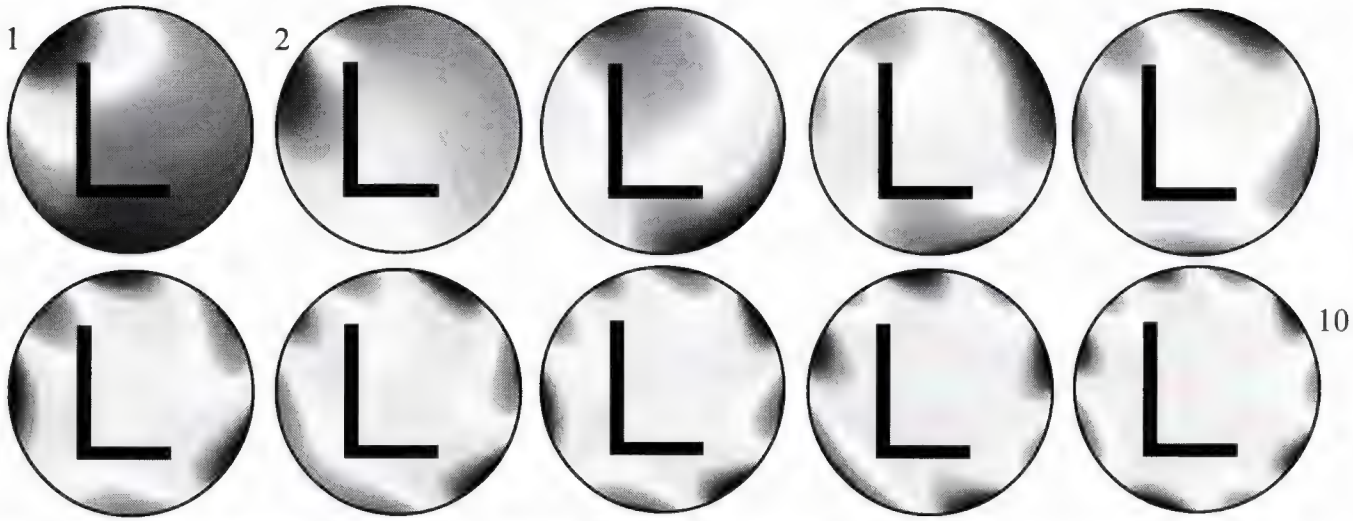


Example





top 10 singular vectors:



From control to estimation

motor systems: generate motor commands that achieve the task goal some time in the future

control $\Rightarrow$ estimation: imagine the future is now and the goal has been achieved, and estimate the motor commands

From control to estimation

motor systems: generate motor commands that achieve the task goal some time in the future

control $\Rightarrow$ estimation: imagine the future is now and the goal has been achieved, and estimate the motor commands

task goals => sensory data

From control to estimation

motor systems: generate motor commands that achieve the task goal some time in the future

control $\Rightarrow$ estimation: imagine the future is now and the goal has been achieved, and estimate the motor commands

task goals => sensory data

motor redundancy corresponds to estimation with incomplete data

From estimation to control

sensory systems: estimate the world state given sensory data

estimation over time: form a prediction/belief about the current state, then correct the belief so as to make it more compatible with the new sensory data

estimation $\Rightarrow$ control: think of the correction as a control signal that pushes the belief towards the sensory data

From estimation to control

sensory systems: estimate the world state given sensory data

estimation over time: form a prediction/belief about the current state, then correct the belief so as to make it more compatible with the new sensory data

estimation $\Rightarrow$ control: think of the correction as a control signal that pushes the belief towards the sensory data

control signals => correction of beliefs

Estimation-control duality in isometric tasks

joint torques ($\mathbf{t}$) = moment arms (M) x muscle activations ($\mathbf{u}$)

control problem:

find $\mathbf{u}$ which minimizes $\frac{1}{2}\left\|\mathbf{t}^* - M\mathbf{u}\right\|^2 + \frac{r}{2}\left\|\mathbf{u}\right\|^2$

estimation problem:

estimate $\mathbf{u}$ given a Gaussian prior $\mathbf{u} \sim N(0,\ I/r)$

and sensory data $\mathbf{y} = \mathbf{t}^*$ assumed to be generated from

$\mathbf{y} = M\mathbf{u} + \varepsilon, \quad \varepsilon \sim N(0,\ I)$

Estimation-control duality in isometric tasks

joint torques ($\mathbf{t}$) = moment arms (M) x muscle activations ($\mathbf{u}$)

control problem:

find $\mathbf{u}$ which minimizes $\frac{1}{2}\left\|\mathbf{t}^* - M\mathbf{u}\right\|^2 + \frac{r}{2}\left\|\mathbf{u}\right\|^2$

estimation problem:

estimate $\mathbf{u}$ given a Gaussian prior $\mathbf{u} \sim N\left(0,\ I/r\right)$

and sensory data $\mathbf{y} = \mathbf{t}^*$ assumed to be generated from

$\mathbf{y} = M\mathbf{u} + \varepsilon, \quad \varepsilon \sim N\left(0,\ I\right)$

the posterior of $\mathbf{u}$ is:

$$p\left(\mathbf{u}|\mathbf{y}\right) \propto p\left(\mathbf{y}|\mathbf{u}\right) p\left(\mathbf{u}\right) = \exp\left(-\frac{1}{2}\left\|\mathbf{t}^* - M\mathbf{u}\right\|^2 - \frac{r}{2}\left\|\mathbf{u}\right\|^2\right)$$

Estimation-control duality in isometric tasks

joint torques ($\mathbf{t}$) = moment arms (M) x muscle activations ($\mathbf{u}$)

control problem:

find $\mathbf{u}$ which minimizes $\frac{1}{2}\left\|\mathbf{t}^* - M\mathbf{u}\right\|^2 + \frac{r}{2}\left\|\mathbf{u}\right\|^2$

estimation problem:

estimate $\mathbf{u}$ given a Gaussian prior $\mathbf{u} \sim N\left(0,\ I/r\right)$

and sensory data $\mathbf{y} = \mathbf{t}^*$ assumed to be generated from

$\mathbf{y} = M\mathbf{u} + \varepsilon, \quad \varepsilon \sim N\left(0,\ I\right)$

the posterior of $\mathbf{u}$ is:

$$p\left(\mathbf{u}|\mathbf{y}\right) \propto p\left(\mathbf{y}|\mathbf{u}\right) p\left(\mathbf{u}\right) = \exp\left(-\frac{1}{2}\left\|\mathbf{t}^* - M\mathbf{u}\right\|^2 - \frac{r}{2}\left\|\mathbf{u}\right\|^2\right)$$

General duality for our class of problems

Written in terms of the desirability function $z(x) = \exp(-v(x))$ the Bellman equation becomes:

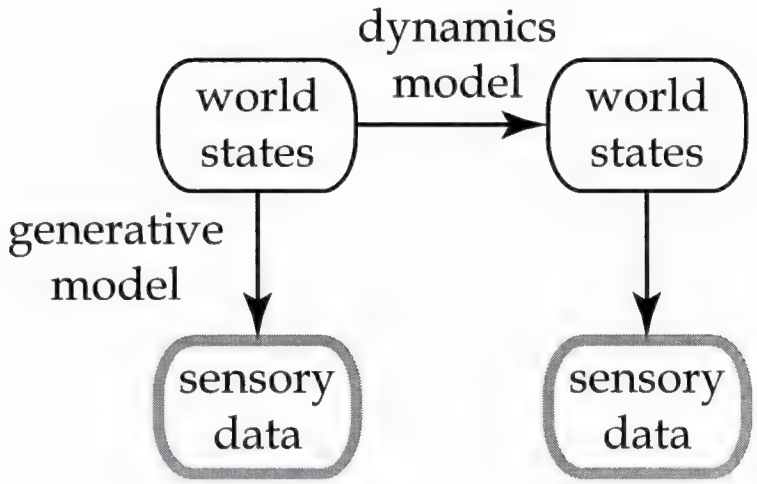
$$z(x_t) = \exp(-q(x_t)) \sum_{x_{t+1}} p(x_{t+1}|x_t)\, z(x_{t+1})$$

This is equivalent to Bayes' rule for backward filtering:

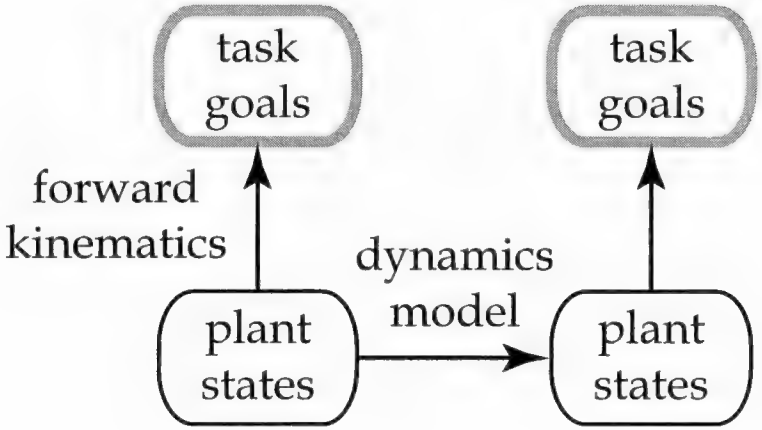
$$b(x_t) = p(y_t|x_t) \sum_{x_{t+1}} p(x_{t+1}|x_t)\, b(x_{t+1})$$

Belief networks for estimation and control

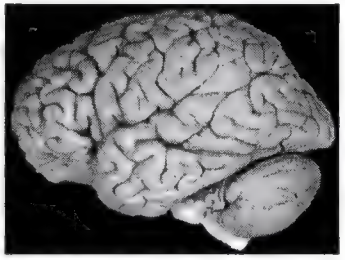
estimation



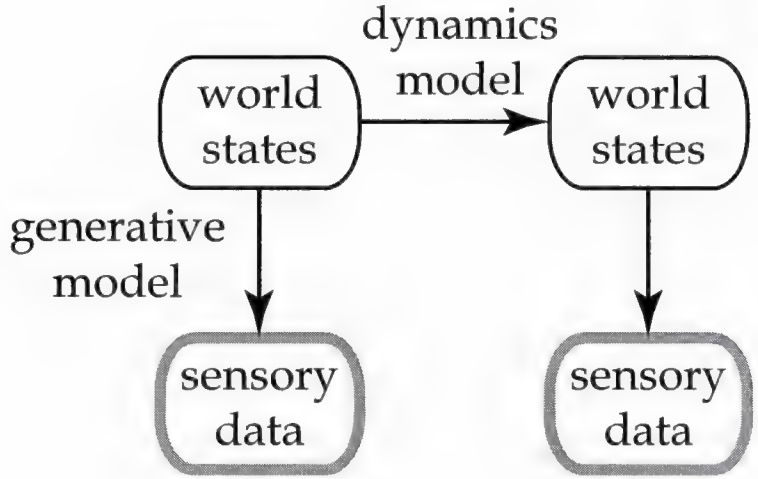
control



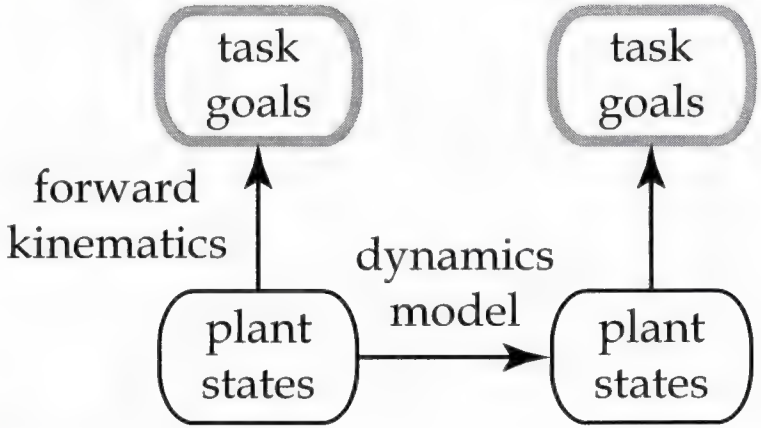
Belief networks for estimation and control

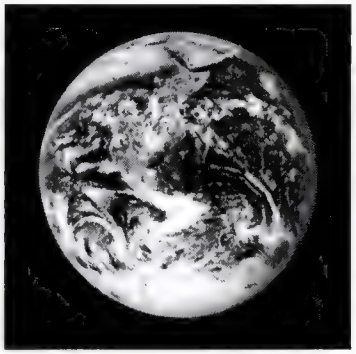


estimation

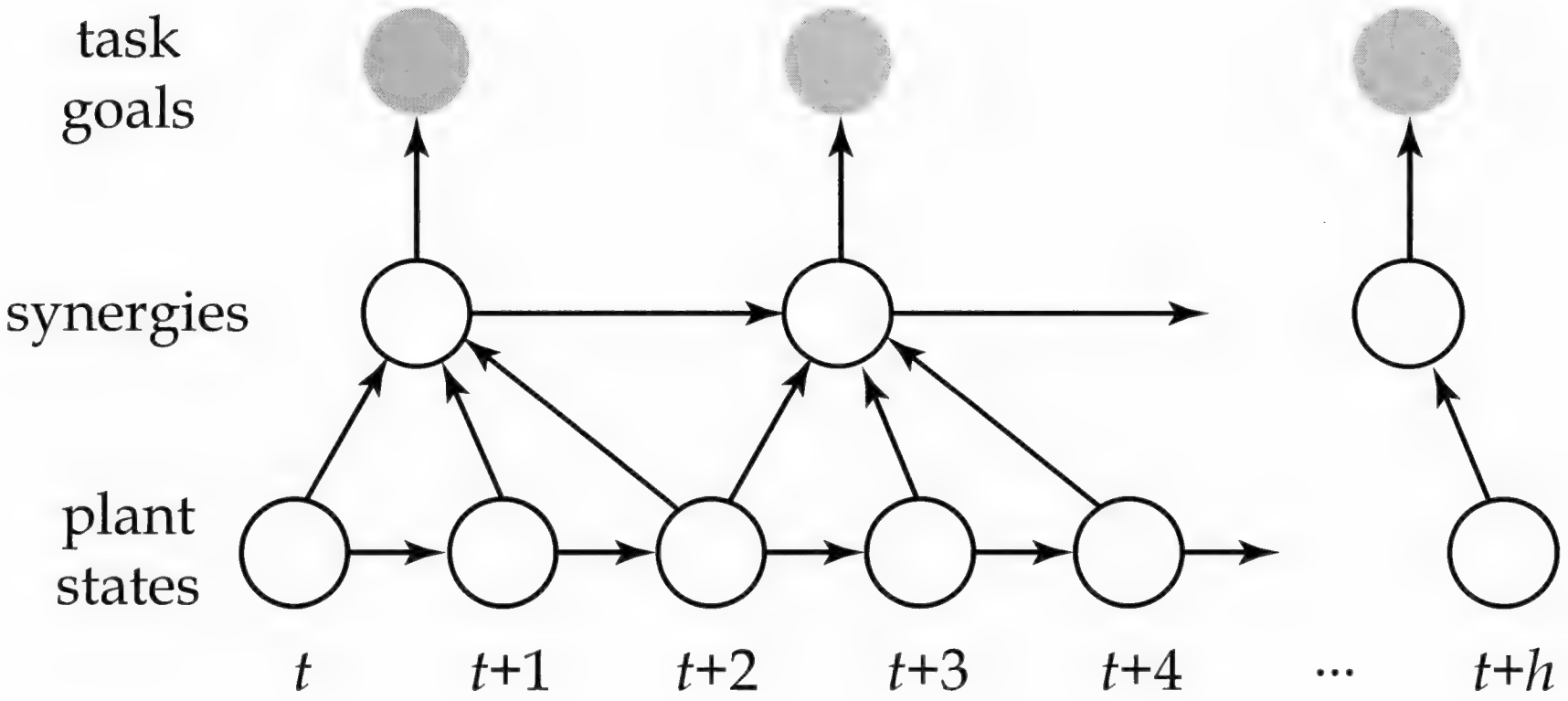


control

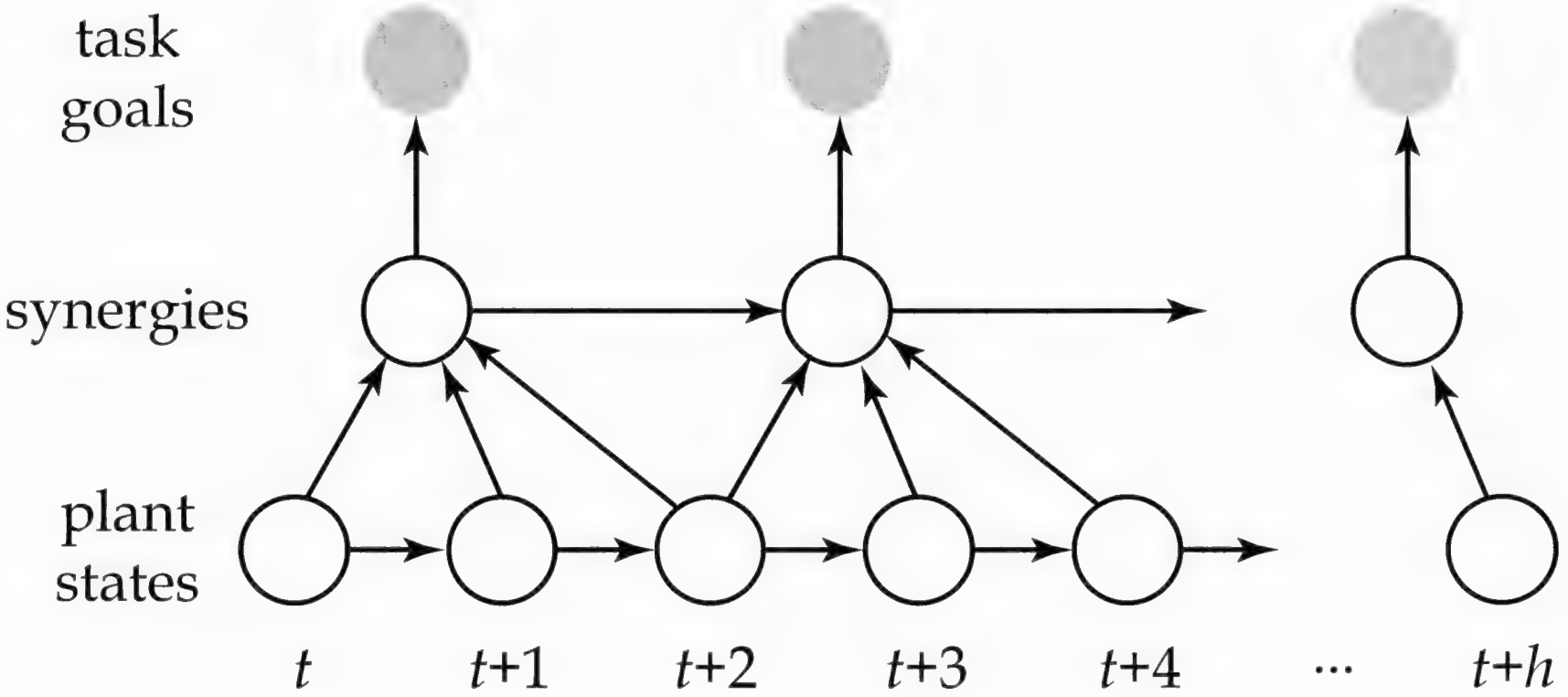




Belief networks for hierarchical control



Belief networks for hierarchical control



motor synergies $\Leftrightarrow$ sensory features

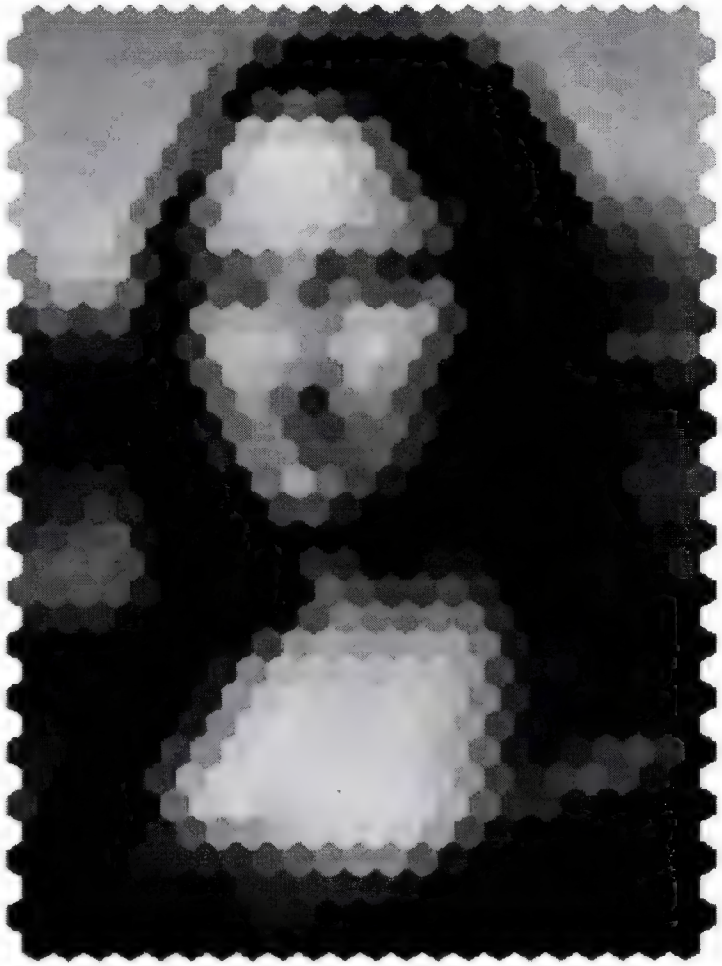
Energy, information, and fly vision

riding the path to discovery on a cycle of experiment and theory

Simon Laughlin

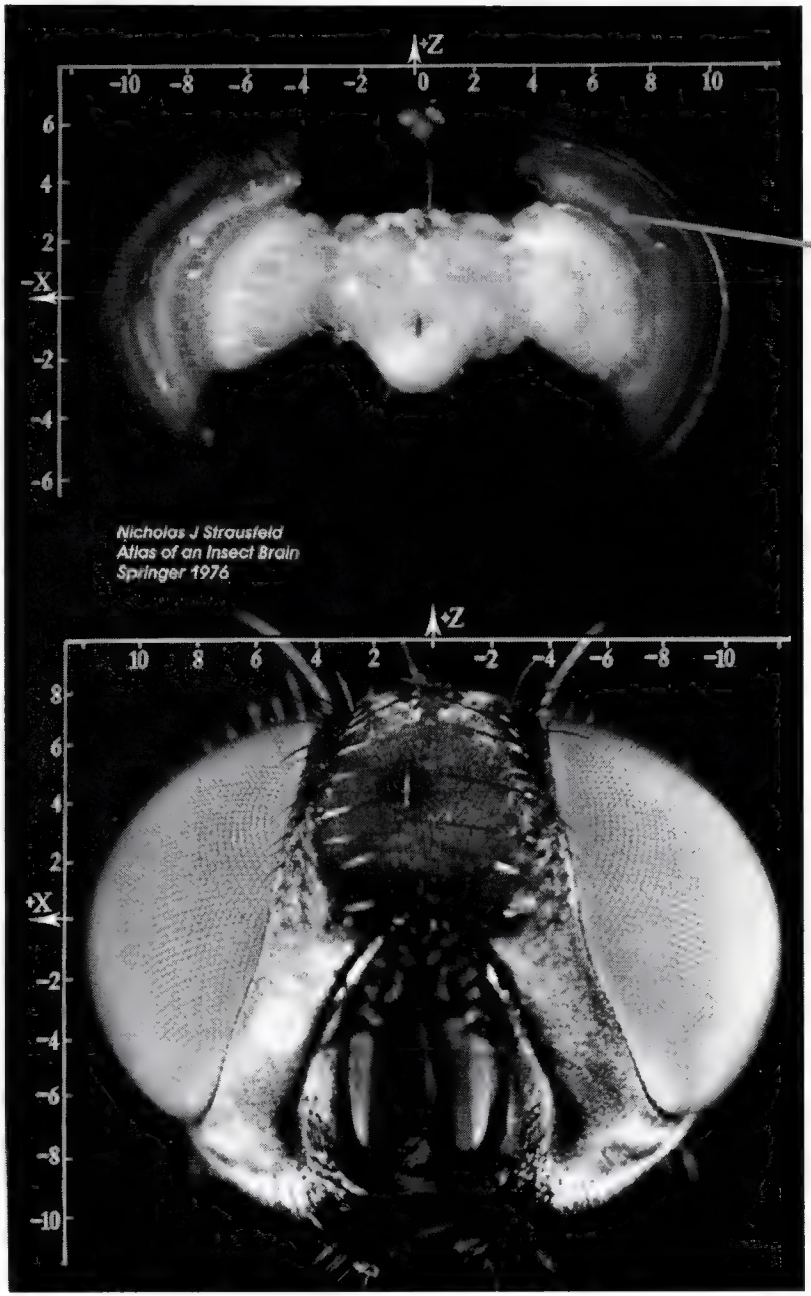
Zoology Cambridge UK

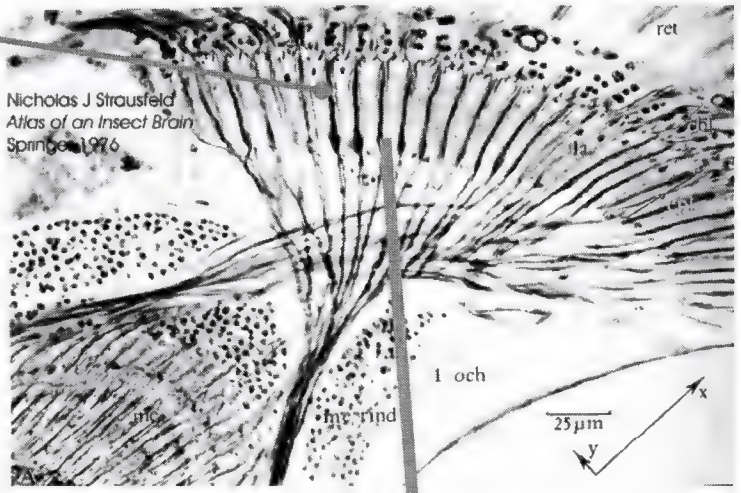
Ommatidia establish the structural unit for neural coding & processing

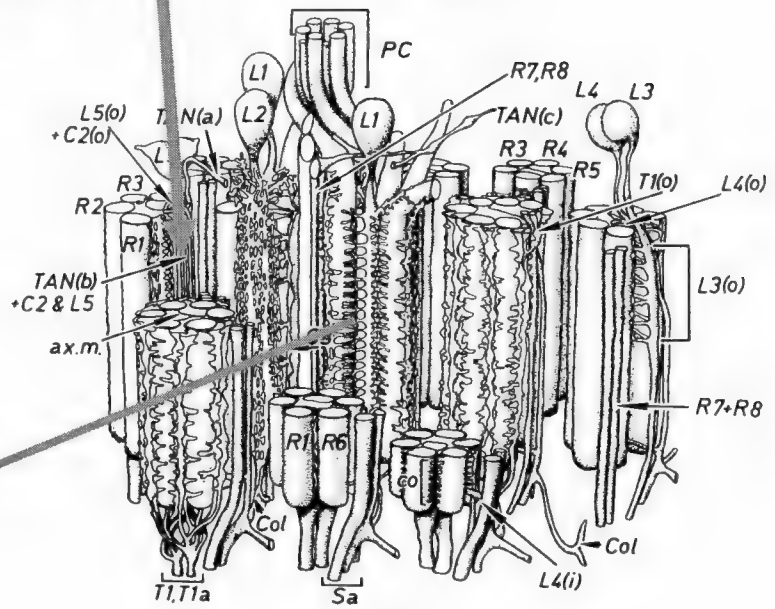


The crystalline neuroanatomy of the fly compound eye,

NJ Strausfeld, 1974 & 1976







LMC

Fig. 1b.

Comprehensive account of neurons and circuit wiring diagram of fly lamina

(compiled by Laughlin Hdbk Sensory Physiology vol VII/6B 1981)

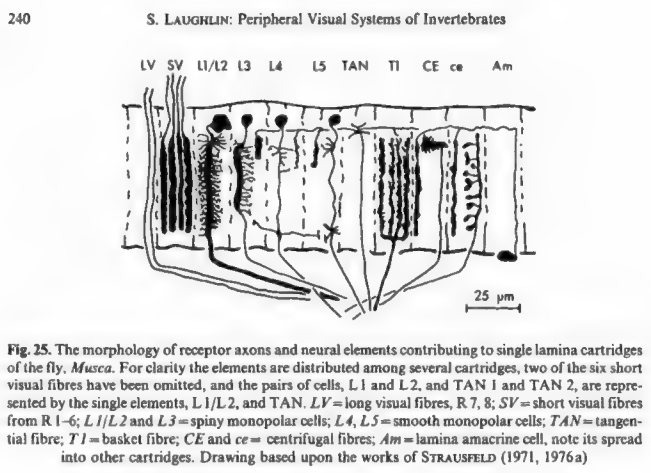


Fig. 25. The morphology of receptor axons and neural elements contributing to single lamina cartridges of the fly, *Musca*. For clarity the elements are distributed among several cartridges, two of the six short visual fibres have been omitted, and the pairs of cells, L1 and L2, and TAN 1 and TAN 2, are represented by the single elements, L1/L2, and TAN. *LV* = long visual fibres, R7, 8; *SV* = short visual fibres from R1–6; *L1/L2* and *L3* = spiny monopolar cells; *L4*, *L5* = smooth monopolar cells; *TAN* = tangential fibre; *T1* = basket fibre; *CE* and *ce* = centrifugal fibres; *Am* = lamina amacrine cell, note its spread into other cartridges. Drawing based upon the works of STRAUSFELD (1971, 1976a)

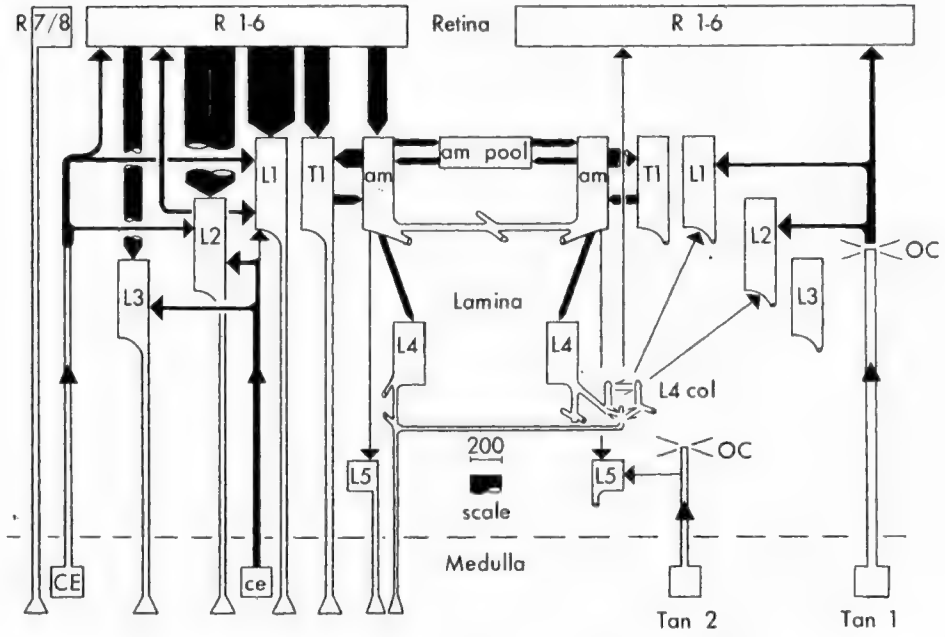
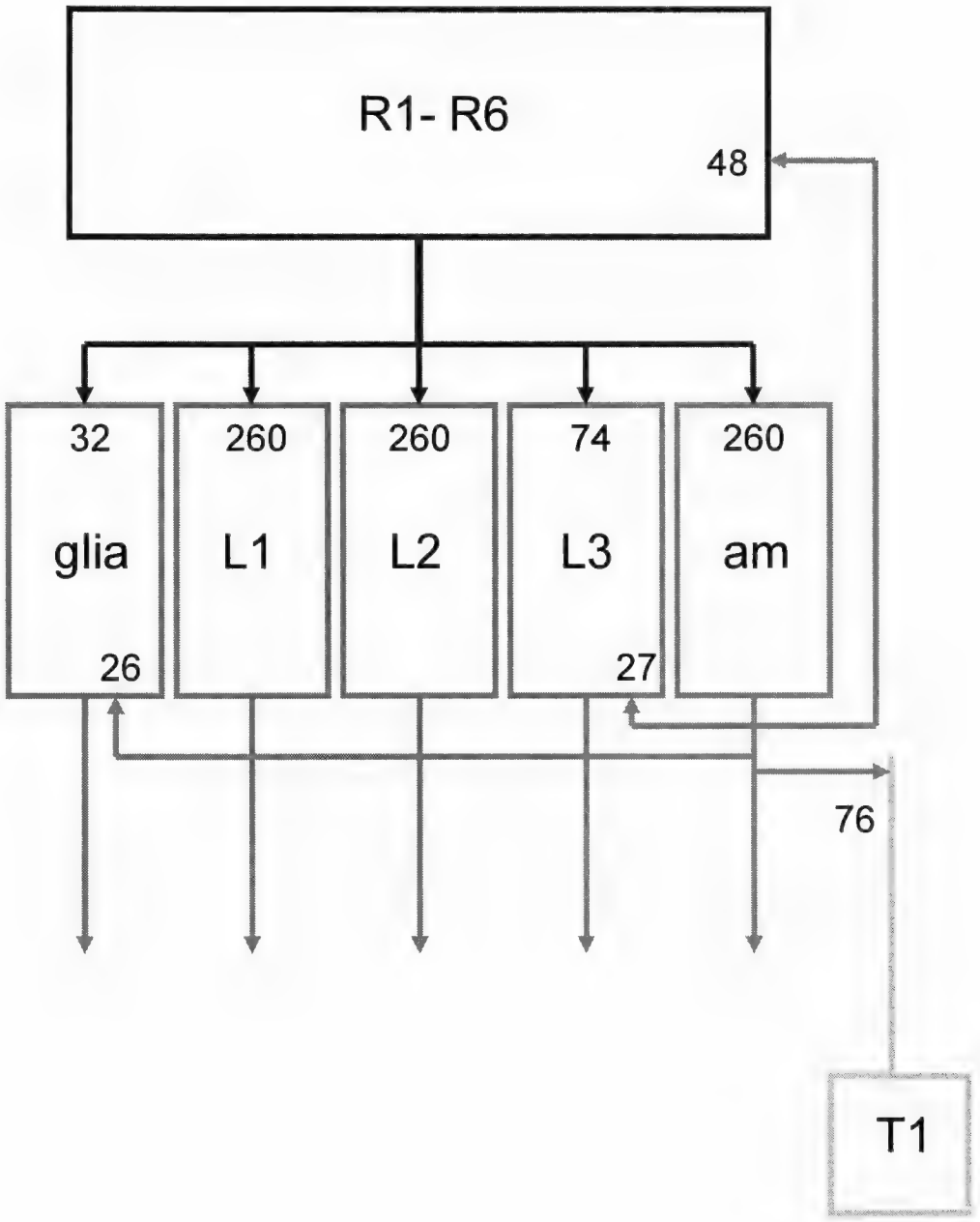


Fig. 26. The anatomical wiring diagram for lamina cartridges of the fly, *Musca*, showing the polarities and frequencies of synaptic contacts between neurons. *Intracartridge* pathways are shown in the left hand cartridge and *intercartridge* connections on the right. The width of *arrows* is proportional to the number of synapses reported by STRAUSFELD (1976a, Tables 7.1 and 7.2) and the scale = 200 contacts. Connections are described in the text. *L4 col* = collaterals from L4 monopolar cells in adjacent cartridges; *OC* = tangential projections to other cartridges

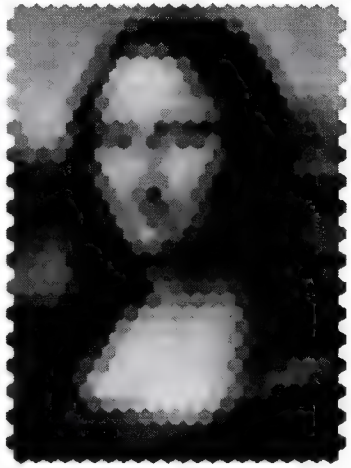


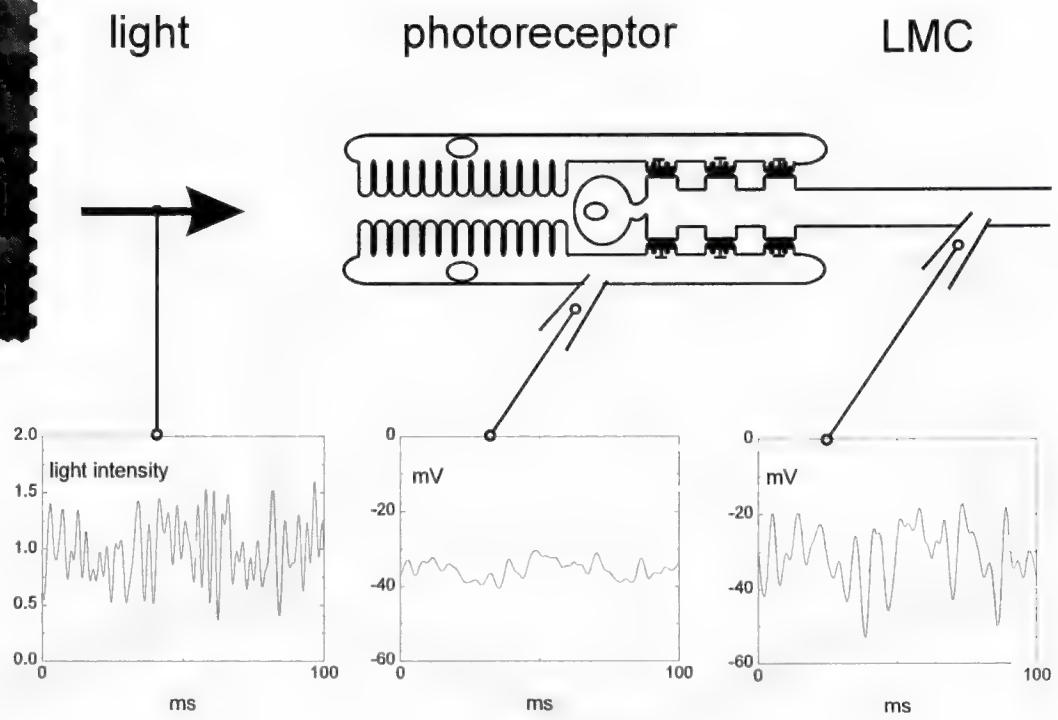
Major elements *Drosophila* lamina

numbers give synapses

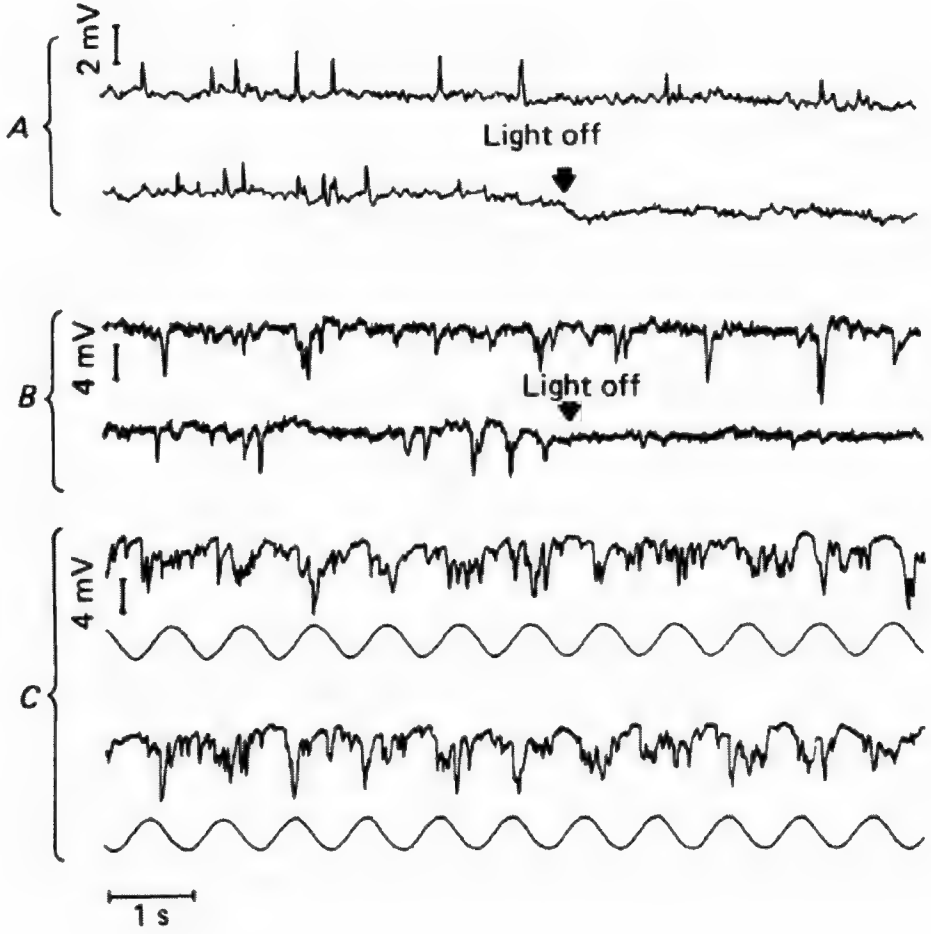
from Meinertzhagen and O'Neil 1991

An elementary circuit in a fly's eye





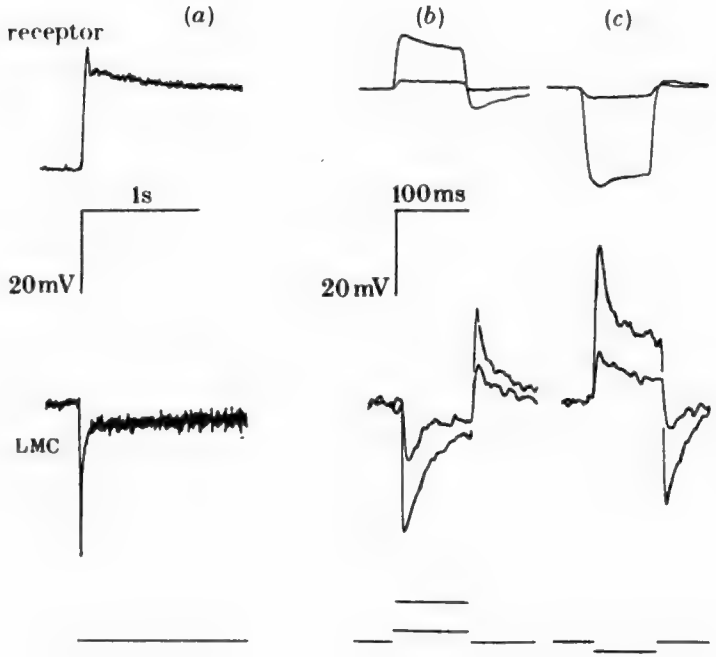
Single photon signals drive fly vision



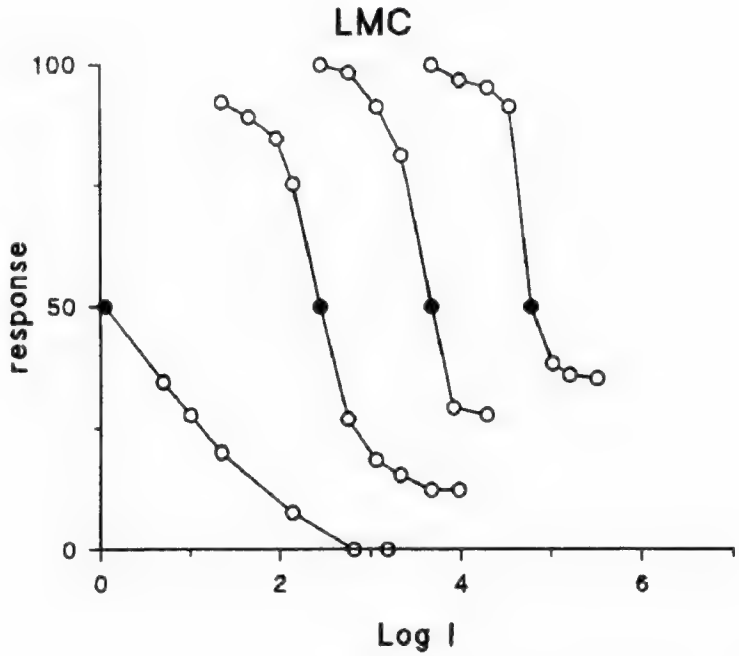
Dubs, Srinivasan and Laughlin, 1981

The responses of blowfly photoreceptors and LMCs, intact retina

Graded responses to changes in intensity



V/log I curves, DA & 3 backgrounds



(Laughlin & Hardie, 1978)

Similar responses and intensity coding properties in the fly compound eye and mudpuppy retina

Laughlin & Hardie 1978

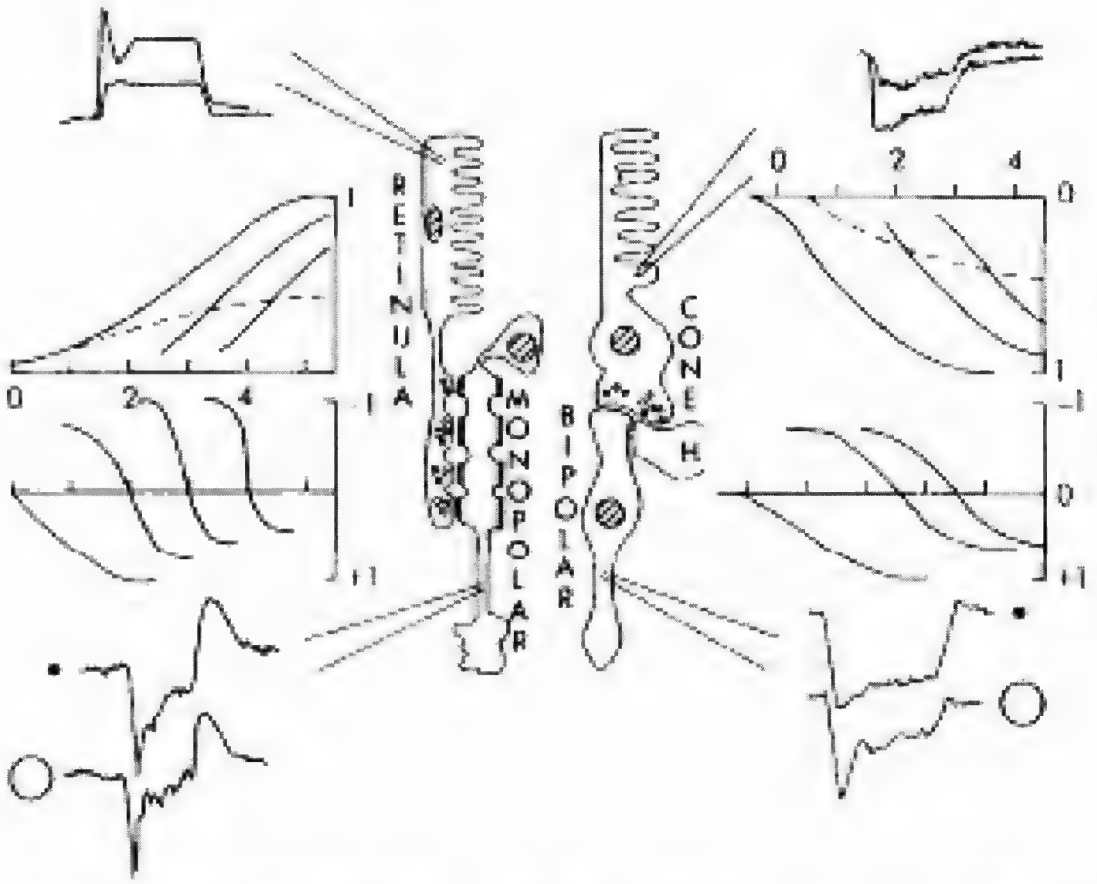


Fig. 11. The similar responses and intensity coding properties of photoreceptors and second order neurons of the compound eye of fly, and the analogous cells in the lens eye of the amphibian *Necturus*. The graded photoreceptor potentials of retinula cells

log – subtract – amplify

effective process for coding contrast (reflectance) over a wide range of mean intensities – shared by insects and vertebrates

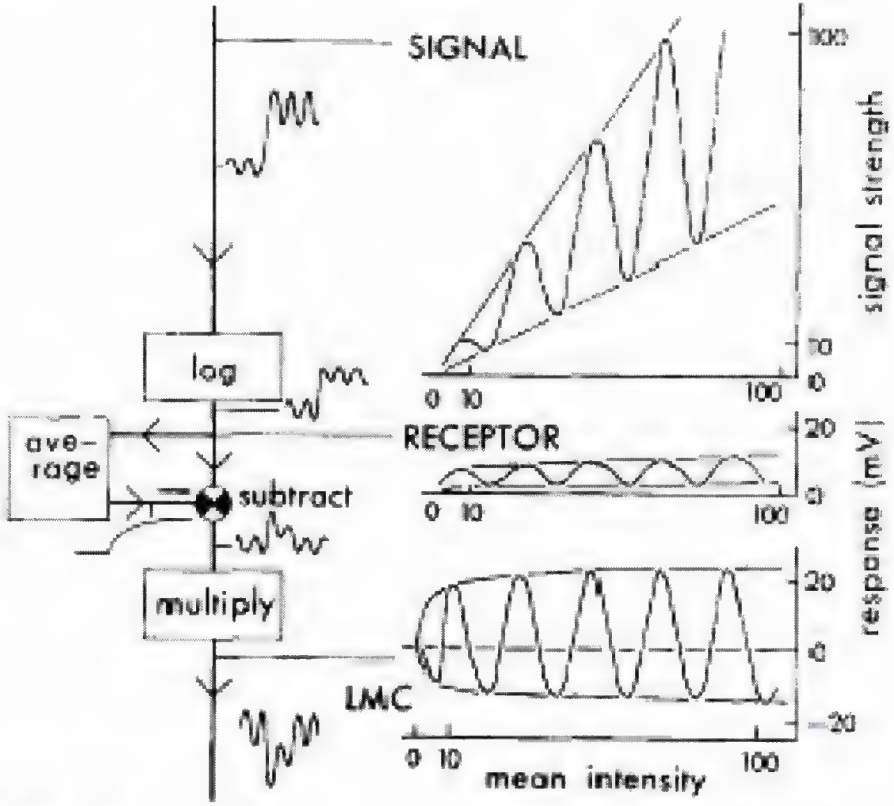


Fig. 12. Diagram summarising the simple 'log transform-subtraction-amplification' strategy for light adaptation as suggested by our data, together with the effectiveness of peripheral adaptation mechanisms in adjusting to the change in input signal strength that accompanies an alteration in mean signal intensity. The block

Laughlin & Hardie (1978)

Design problem, how best to implement *log – subtract - amplify*

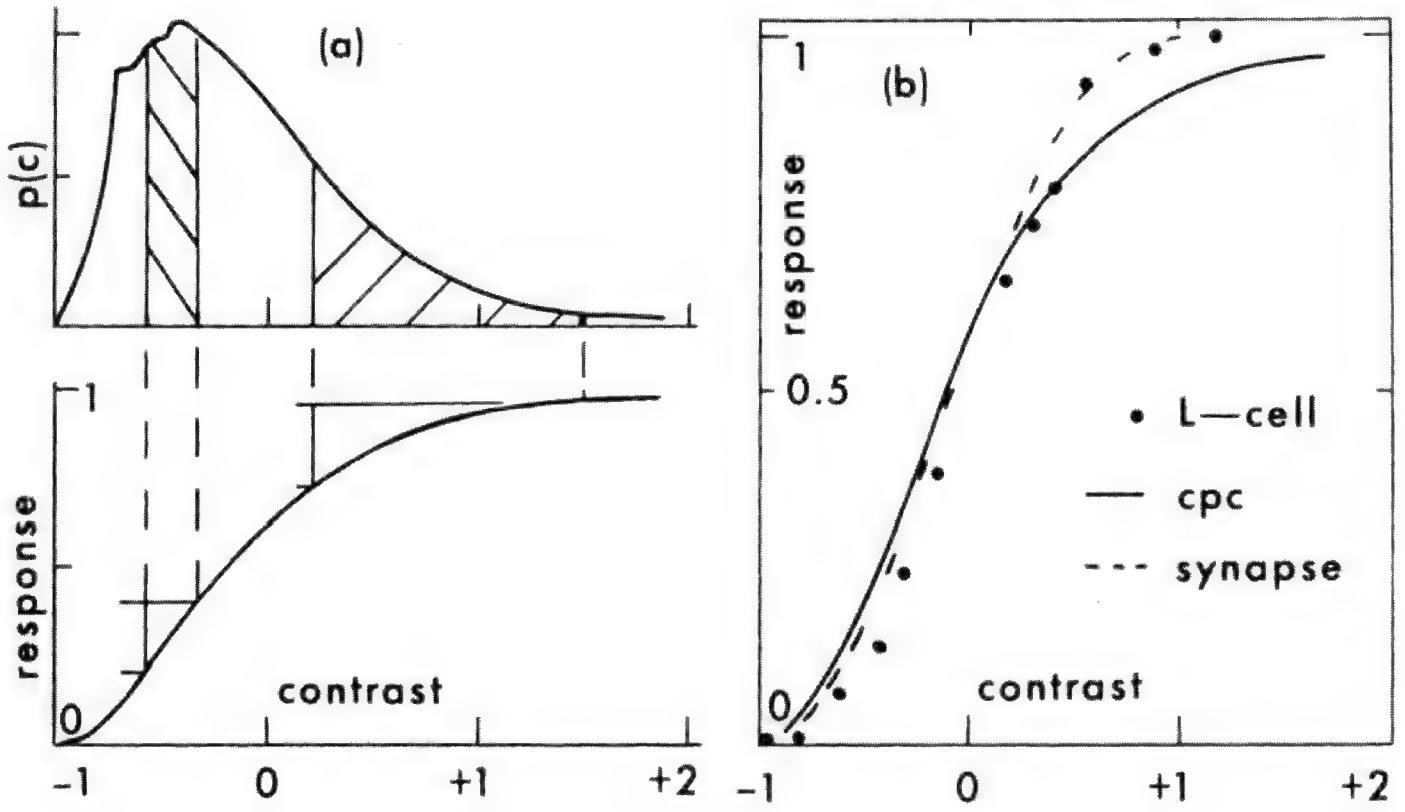
Processes

- What do you subtract?
  - Global mean vs. local mean
  - If local, how local?
- What gain do you use for amplification?

Measurements → observations → questions

Information **theory** predicts that the optimum gain should follow an input statistic, the cumulative probability distribution of amplitudes, to ensure that all neural response levels are used equally often

experiment



(Laughlin, 1981)

Transforming signals from olfactory receptor neurons to projection neurons in *Drosophila* olfactory glomeruli

Bhandawat et al 2007

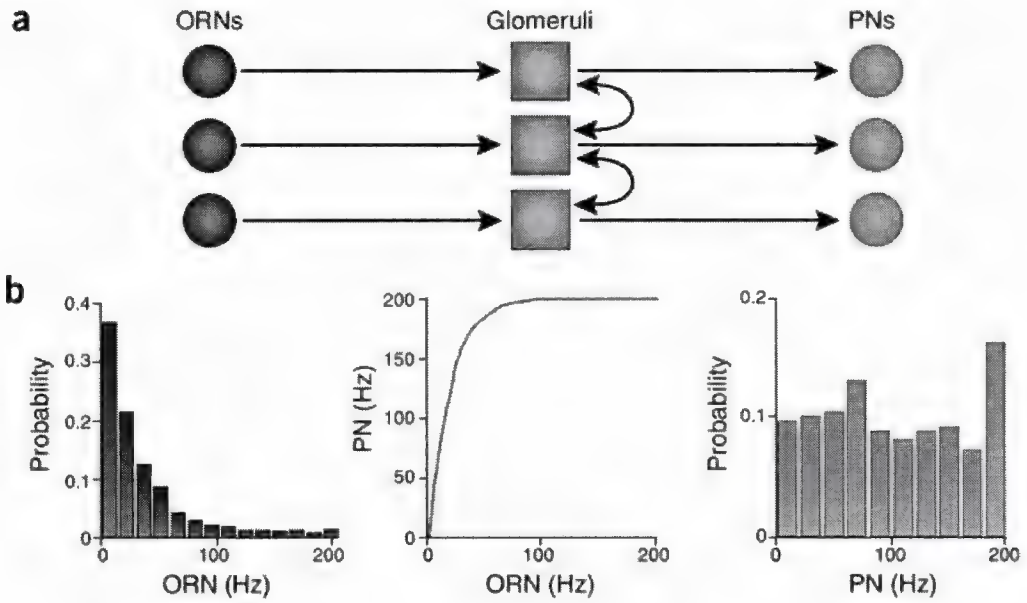
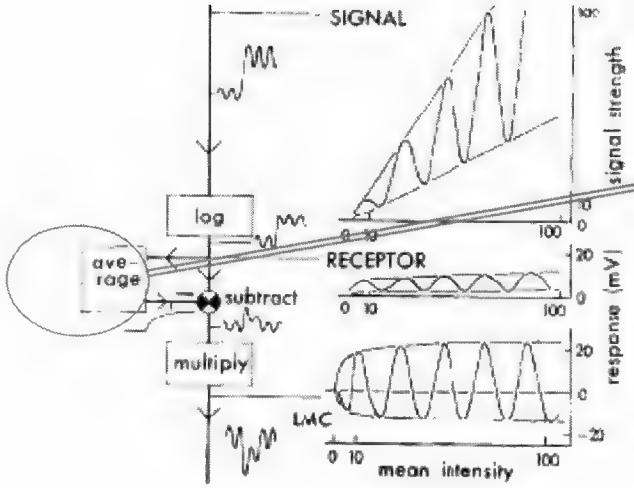


Figure from Abbott & Luo,
Nature Neurosci News and Views 2007

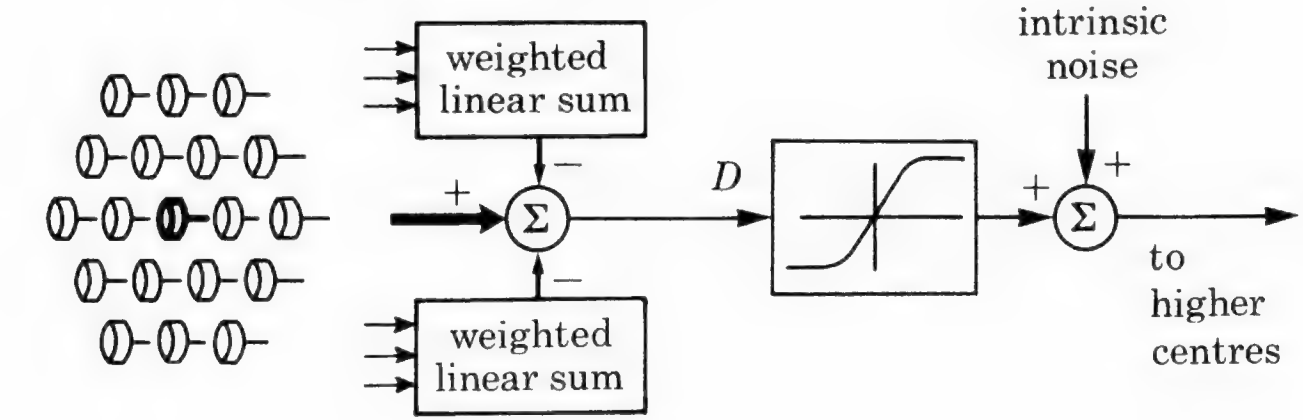


How do you get the average that gives you the mean to subtract?

exploit fact that elements within pictures are not randomly distributed, which means that pixels are correlated, strongest with neighbours

Predictive coding

Srinivasan et al 1982



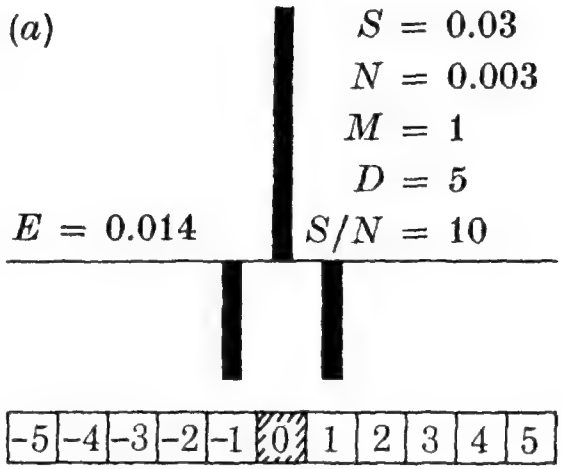
THEORY

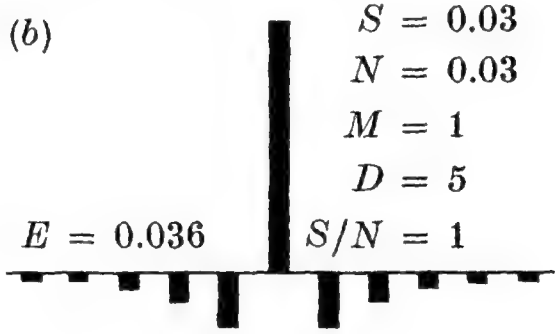
Predictive coding requires an intensity dependent surround

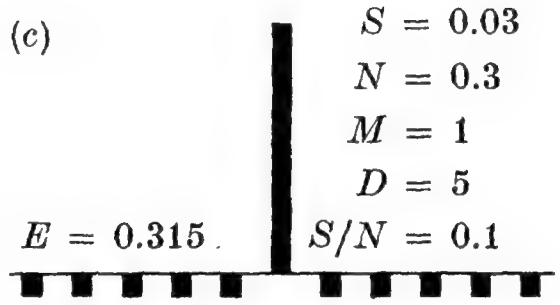
PREDICTION

Take wider samples when input is unreliable (photon noise)

Experiment

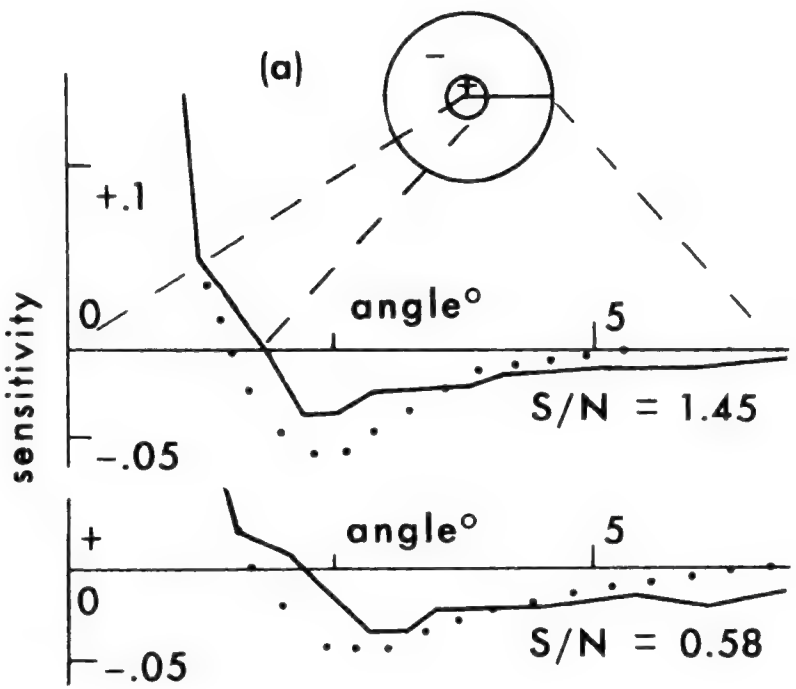


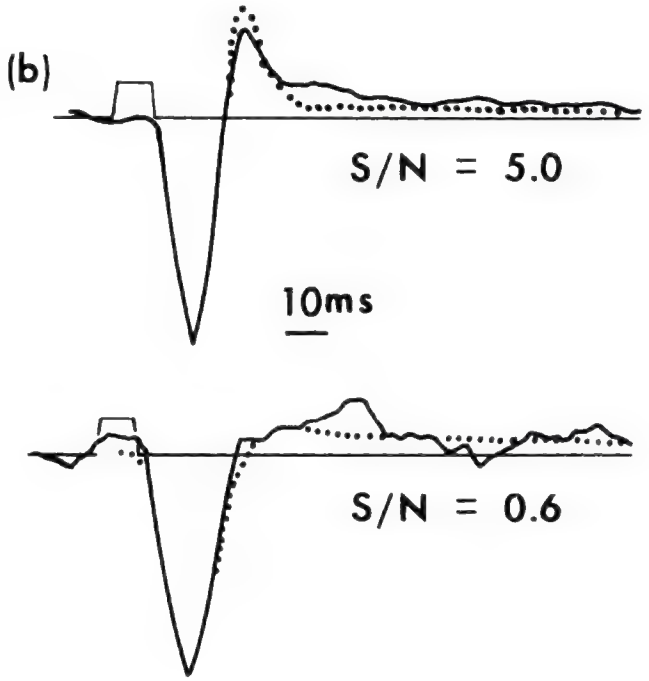




Experiments testing for predictive coding

(Srinivasan, Dubs and Laughlin, 1982)



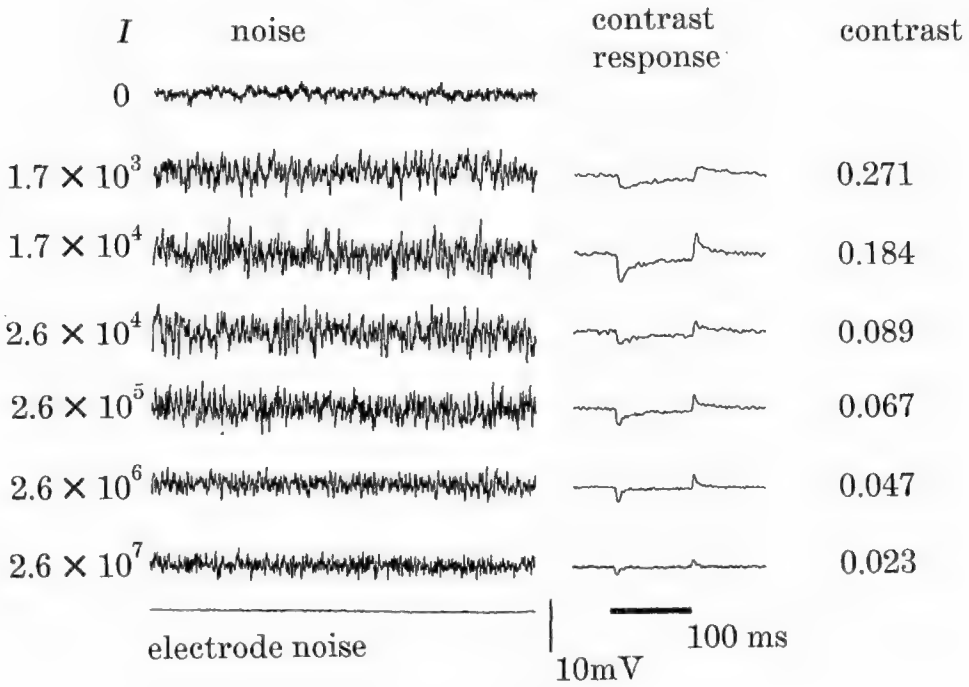


——— Experimental data, from typical cell

····· Theoretical prediction

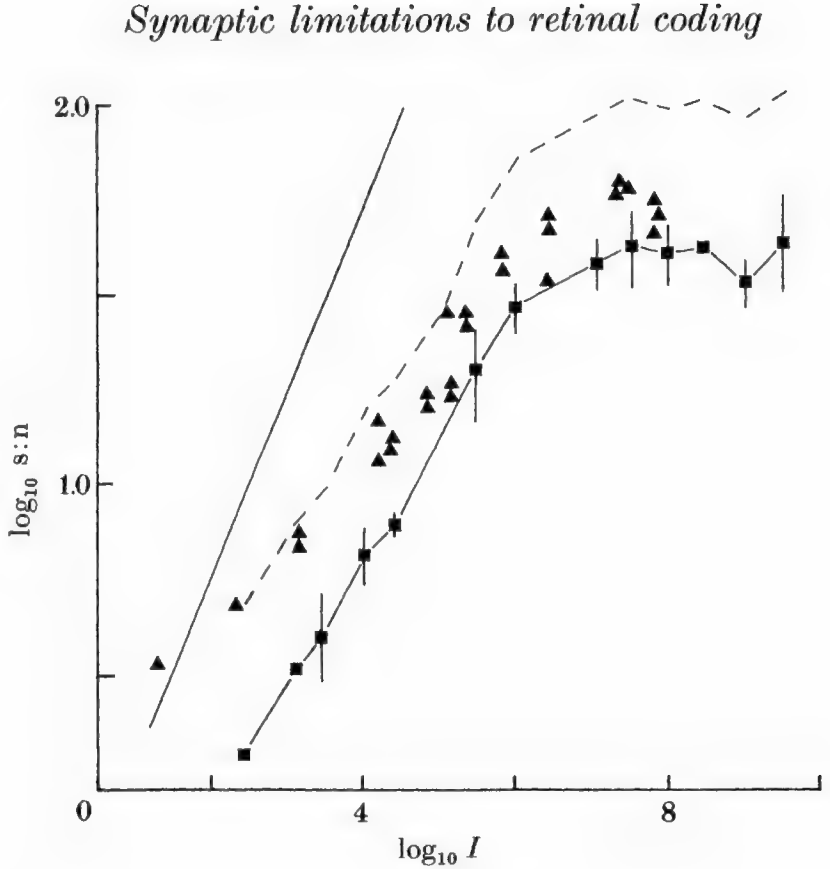
N.B. S/N was measured *in situ* for each LMC

Requirement of theory
signal encounters significant noise

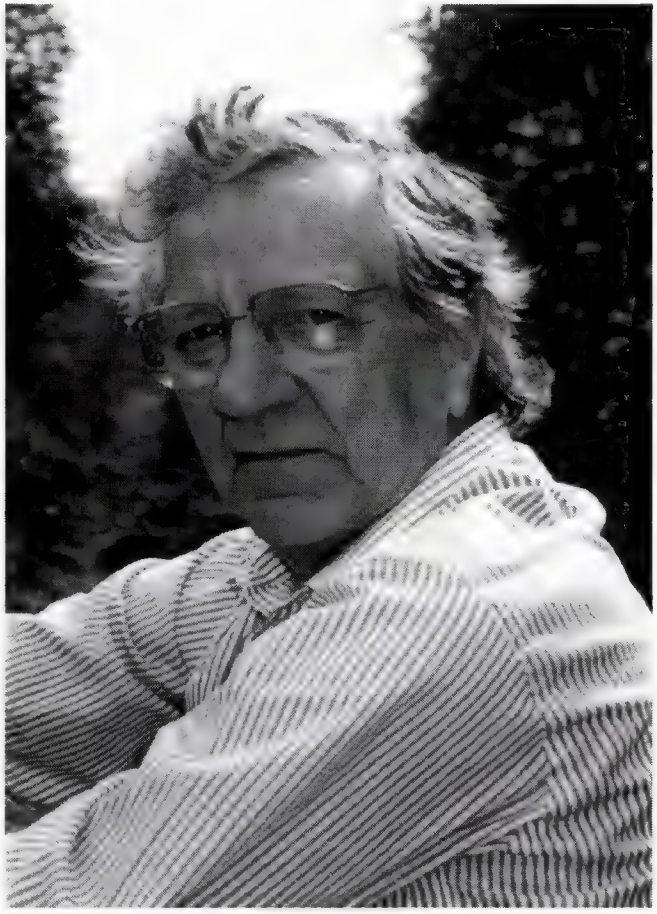


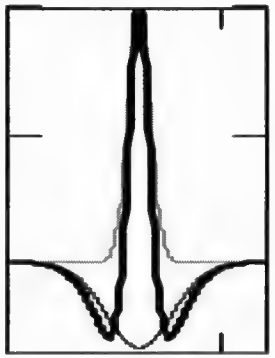
Experiments Photoreceptor and LMC SNRs at different background light levels, given in effective photons/photoreceptor/s

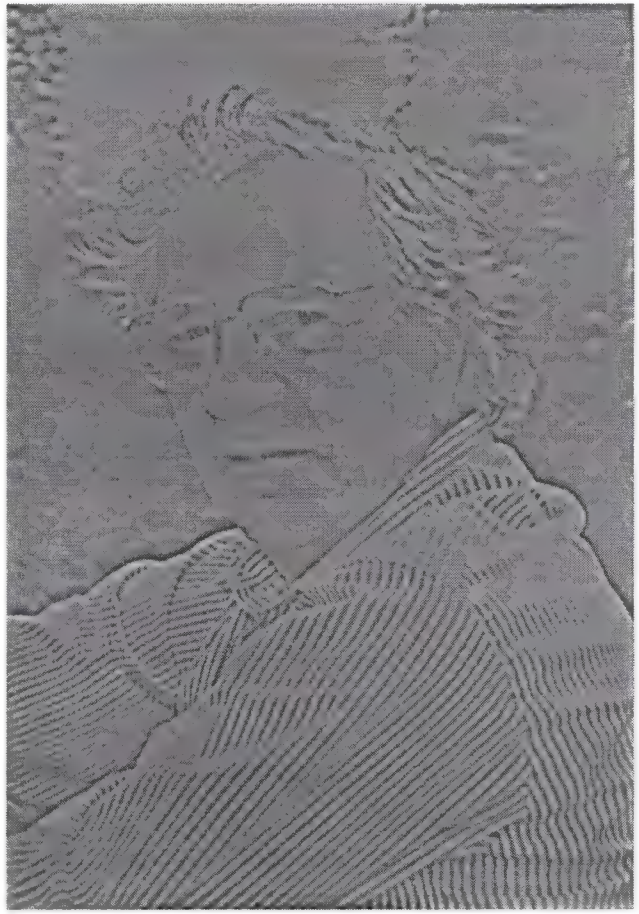
Laughlin, Howard & Blakeslee, 1987



“The neat packaging of information”
Horace Barlow 1961
redundancy reduction as a goal of early sensory processing



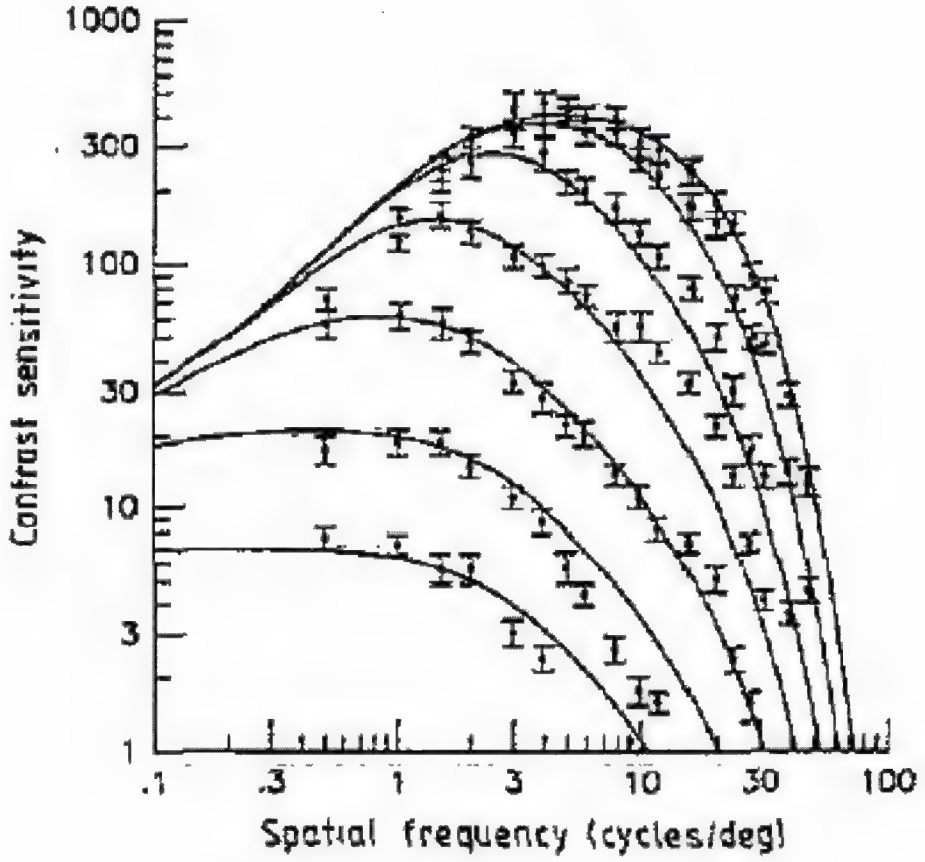




Conclusion from experiments on predictive coding

- Spatial and temporal antagonism designed to "package neatly", ***eliminate redundancy***
- Redundancy reduction is intensity dependent
- Effective processing requires ***adaptation*** – input dependent changes in processing by the circuit
- Applies to human vision
- Explains why vision changes from low-pass to band-pass as intensity increases, and how much it changes

Similar redundancy reduction model applied to human vision (Atick & Redlich, in Atick, 1992)



Data points = psychophysical data; Curves = redundancy reduction model

GREAT!

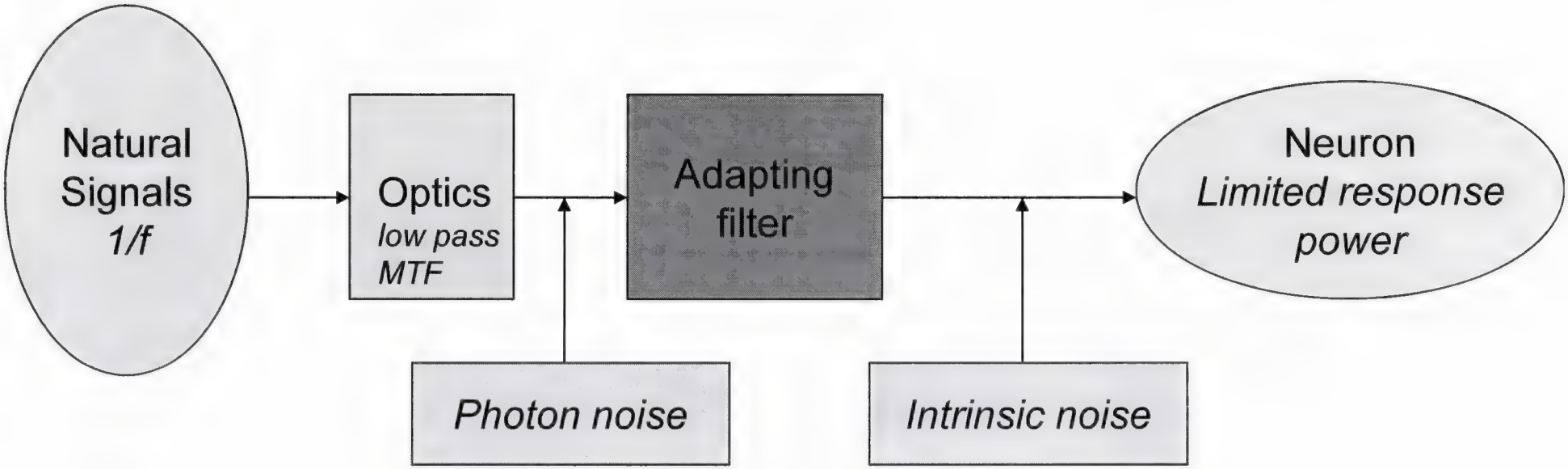
NO

Theory inadequate

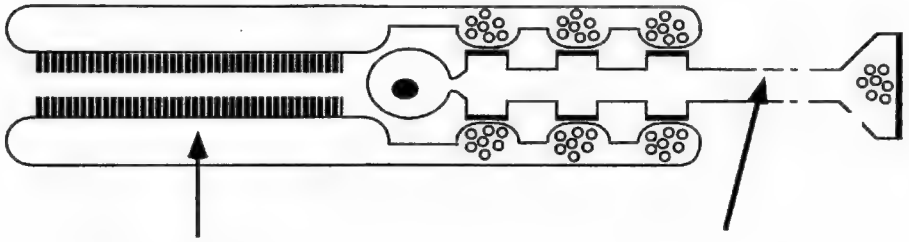
A filter that maximises information rate within constraints on signal and noise
(Hans van Hateren, 1992, *J comp PhysiolA*)

Applied information theory to signals with natural statistics

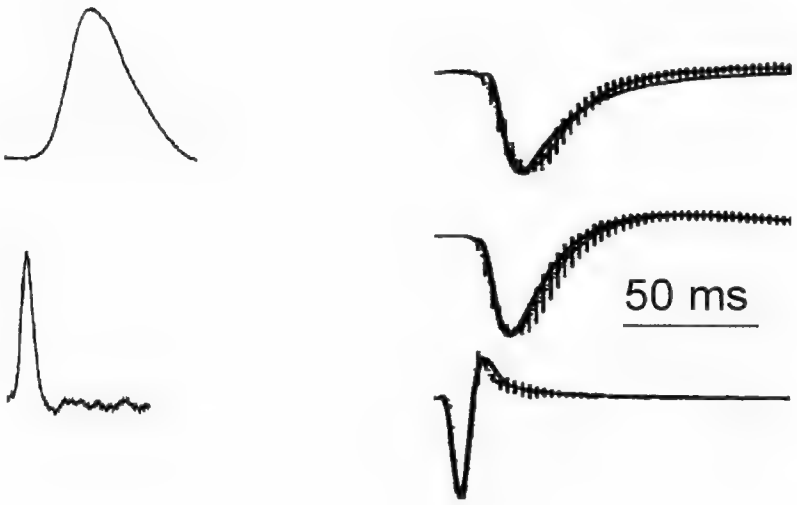
derived filter that maximises information carried by a neuron with a limited power output



Fit of van Hateren's optimum filter model to his LMC data at 3 light levels



R1-6 LMC



LMC Data |

Model ——

Very dim

Quite dim

daylight

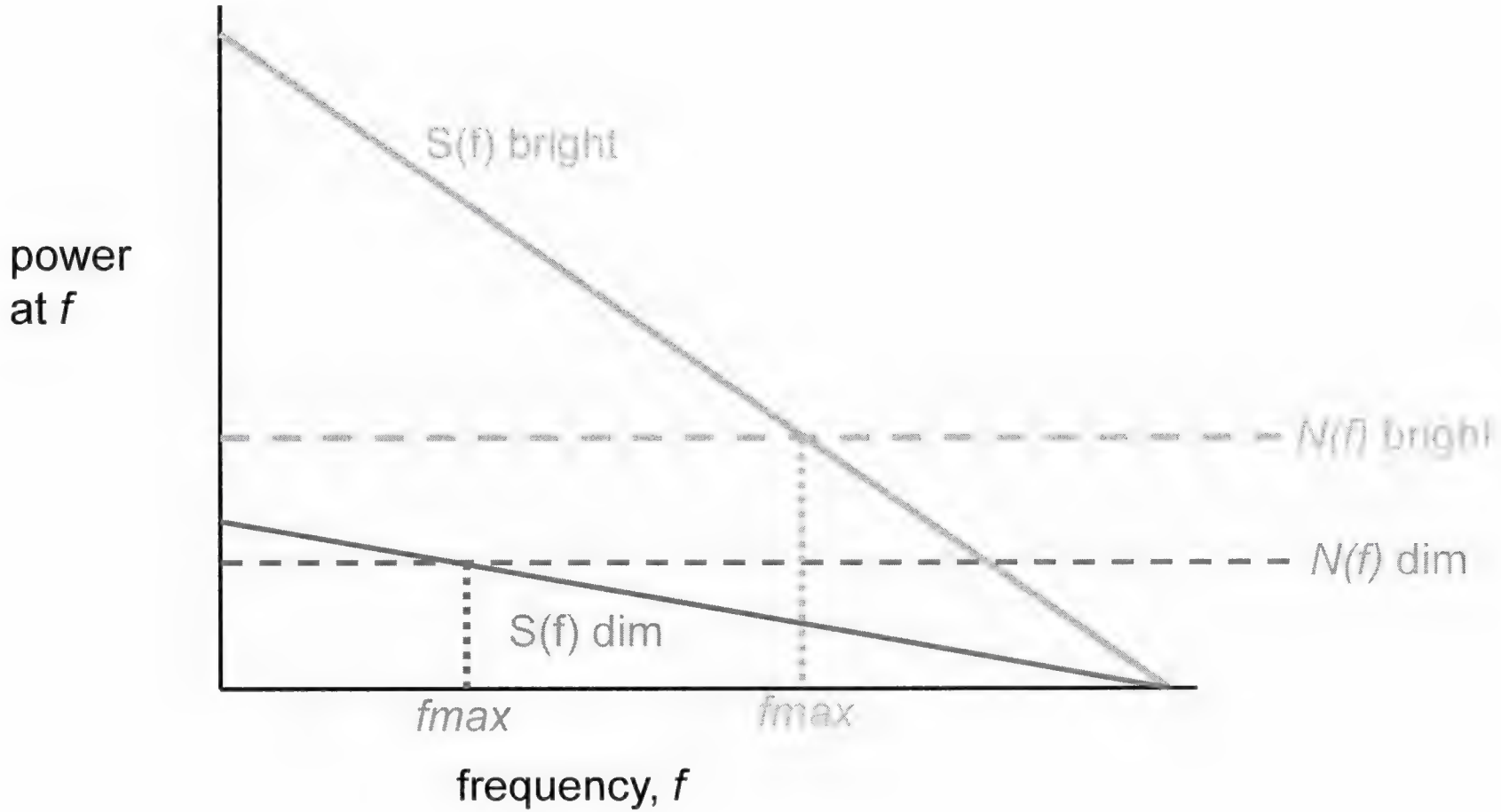
Versatility
Adaptability
Adaptation

Transfer function matched at two successive levels

Phototransduction Cascade

Synapse

When the signal spectrum falls off with frequency the highest useful frequency, *fmax*, decreases with light level, because the spectrum of photon noise is flat and the S/N follows the square root law



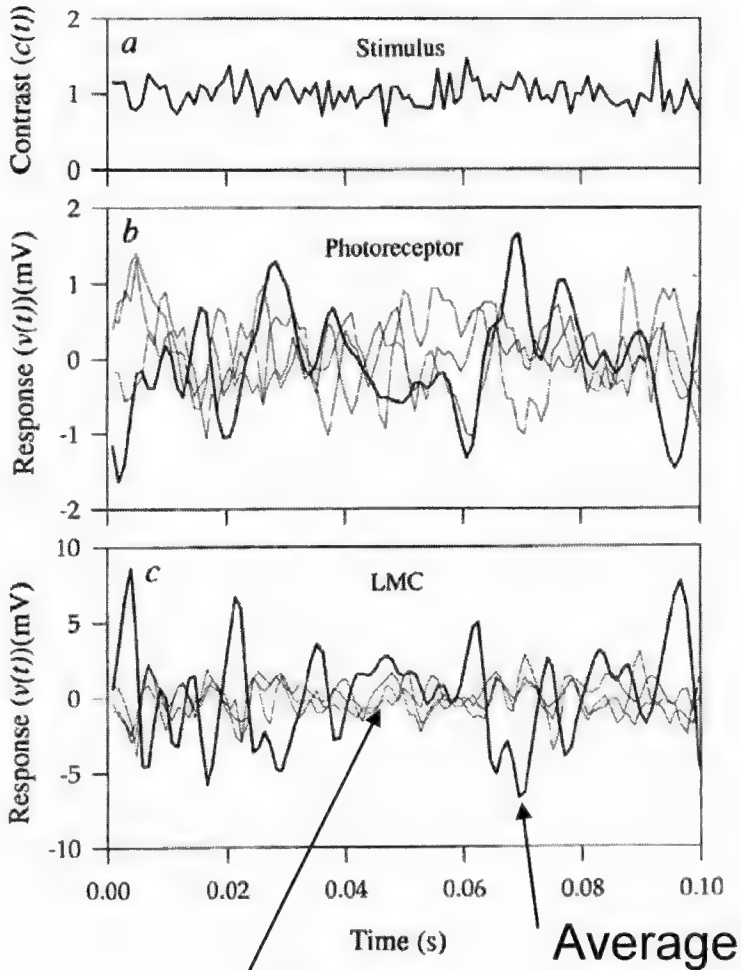
Why has the cycle gone this far?

- Good data
  - superlative anatomy
  - high quality recording,
  - Identified cells in circuits
- Tell us what's going, pose right questions
- Good theorists
- Develop theory that answers question
- Validated by testable predictions

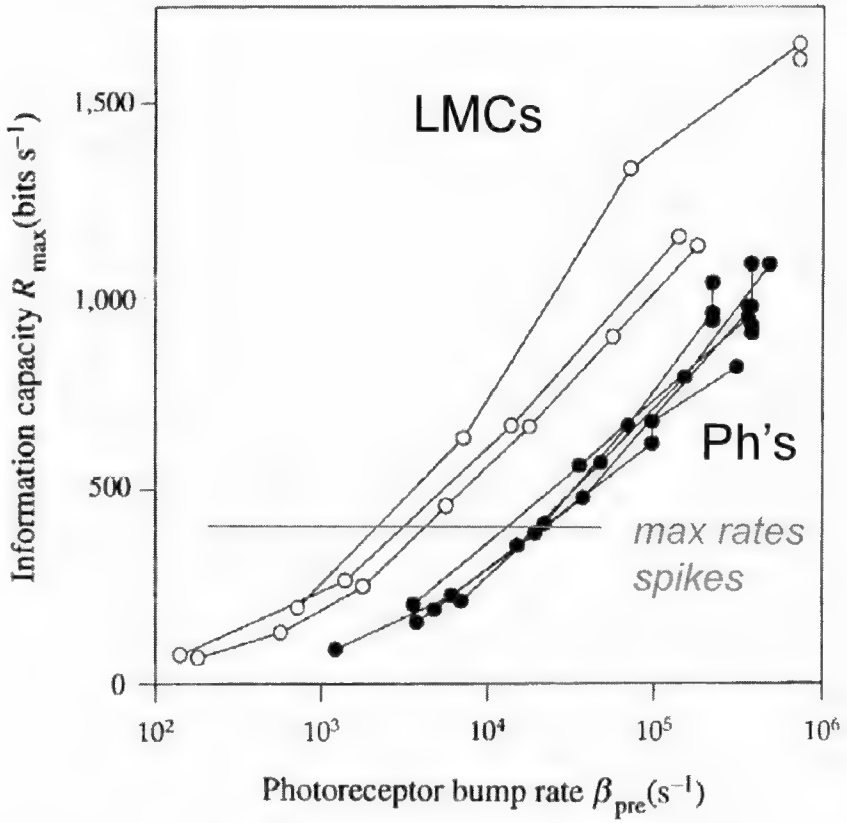
Loosing the plot?

- More questions
- Coding optimised to transfer information within constraints of noise and response range
- How much information does an optimally coding neuron transmit?

Measuring information capacities in blowfly photoreceptors and LMCs



de Ruyter van Steveninck & Laughlin, 1996

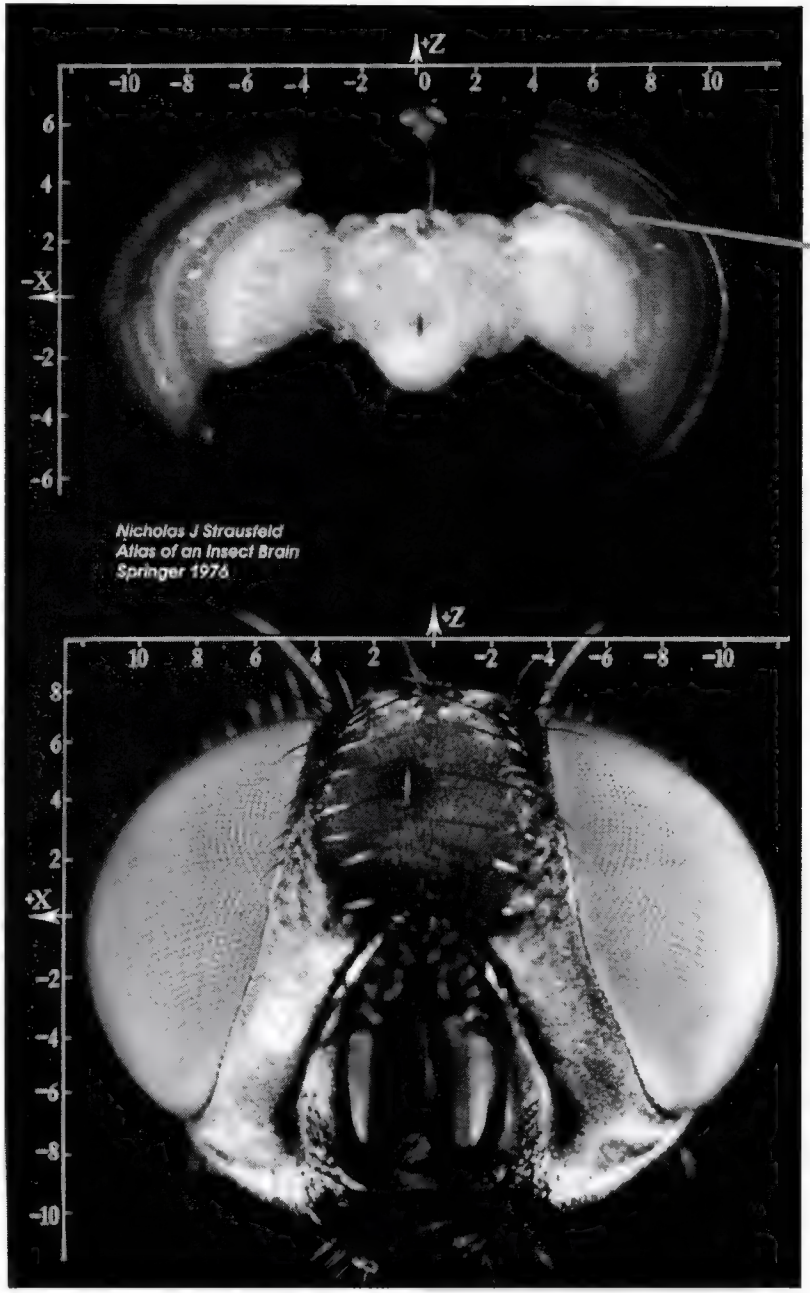


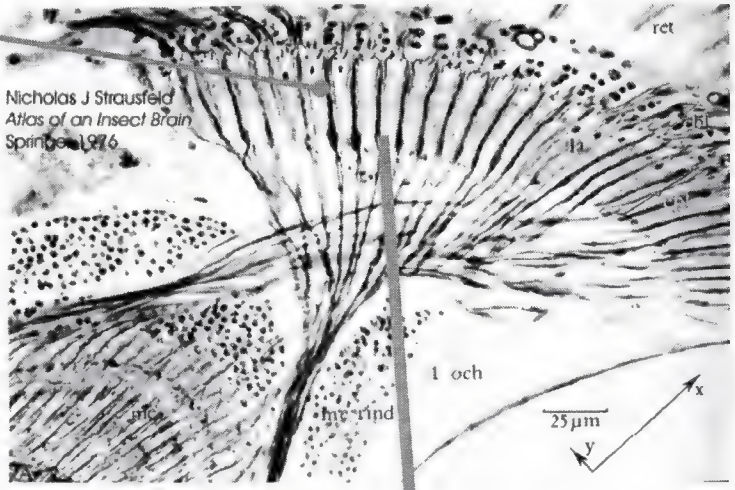
Average response over many presentations = *signal*

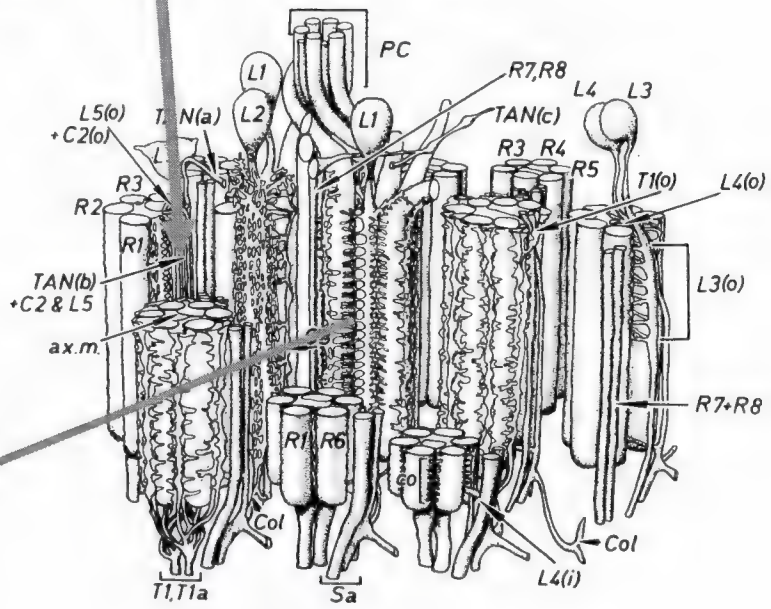
Deviations from av. on individual presentations = *noise*

The crystalline neuroanatomy of the fly compound eye,

NJ Strausfeld, 1974 & 1976







LMC

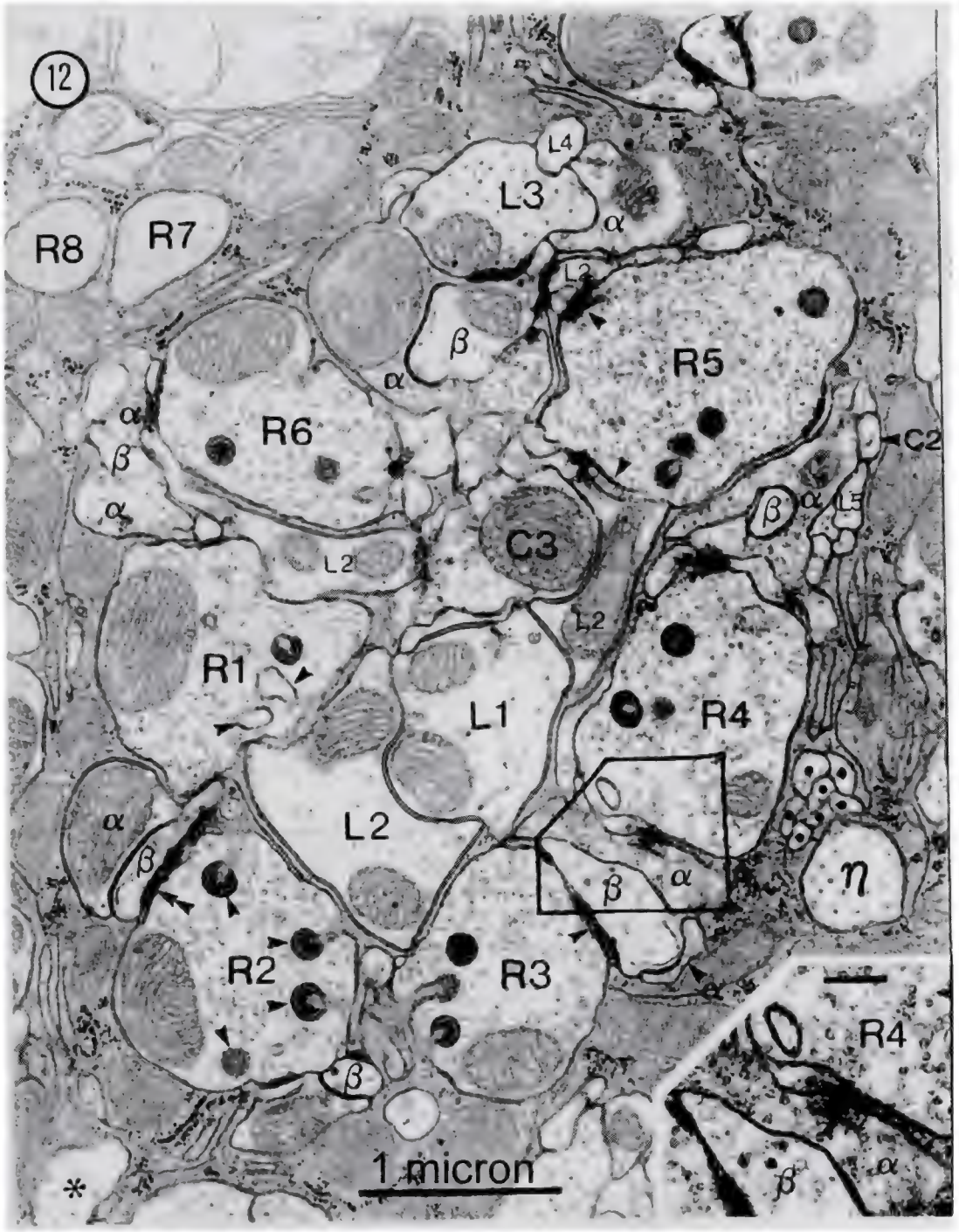
Fig. 1b.

EM section of the type used to make complete serial reconstructions of lamina cartridges in flies

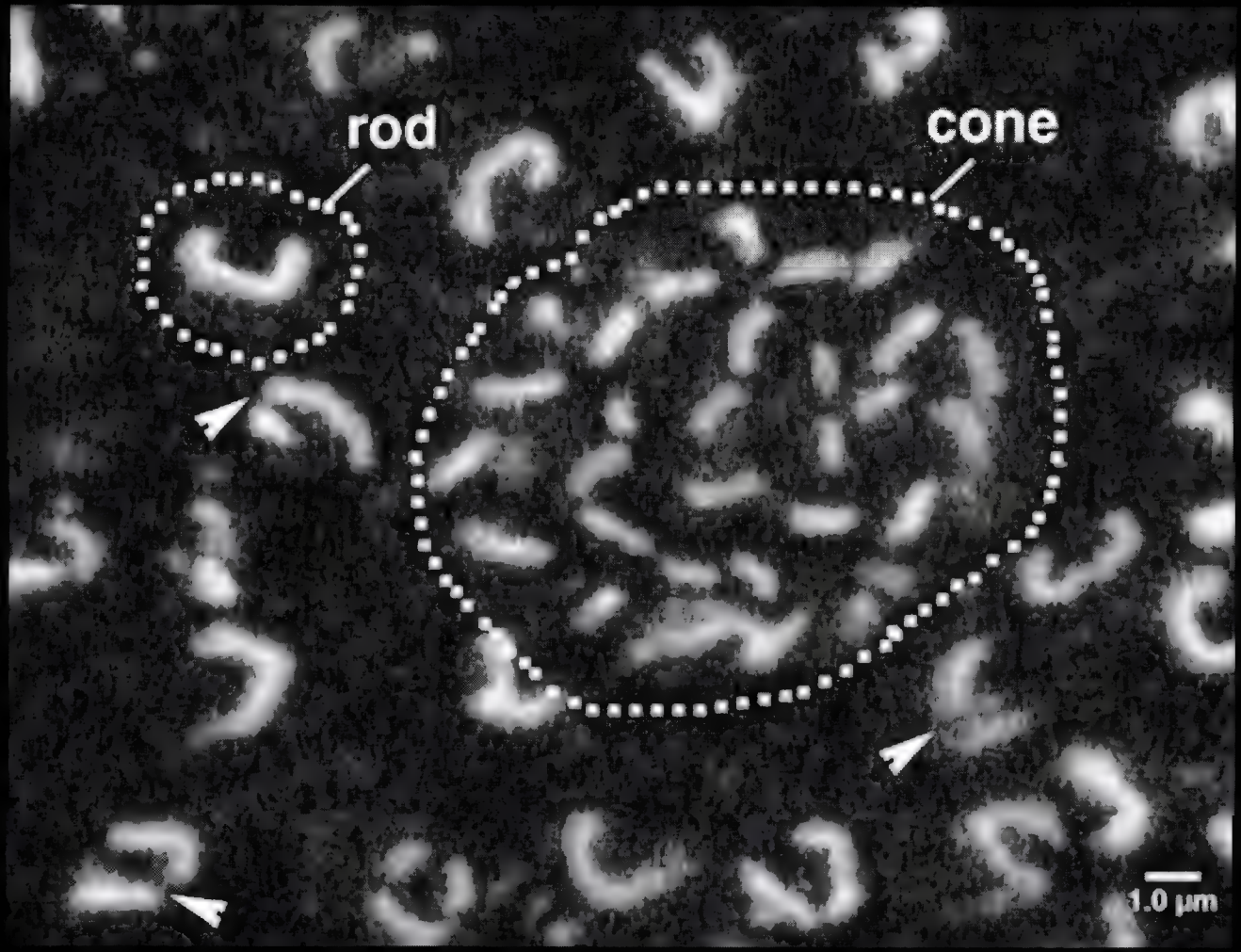
Meinertzhagen Lab, Dalhousie

www.flybrain.org

1320 synapses onto each LMC



cone terminal has 20x more active zones



Peter Sterling data, from Wässle et al., 2003

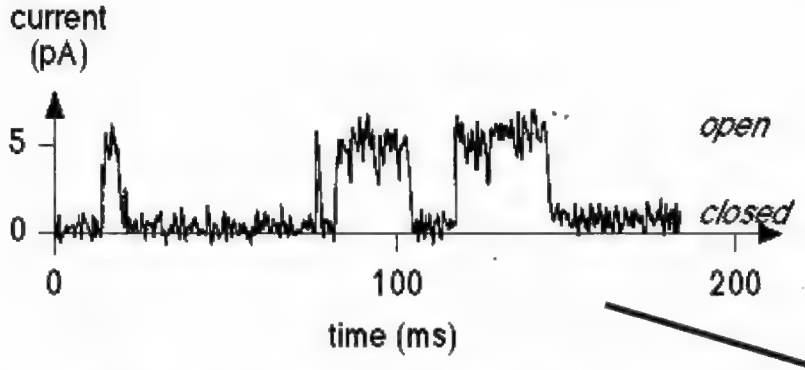
macaque

Conclusion

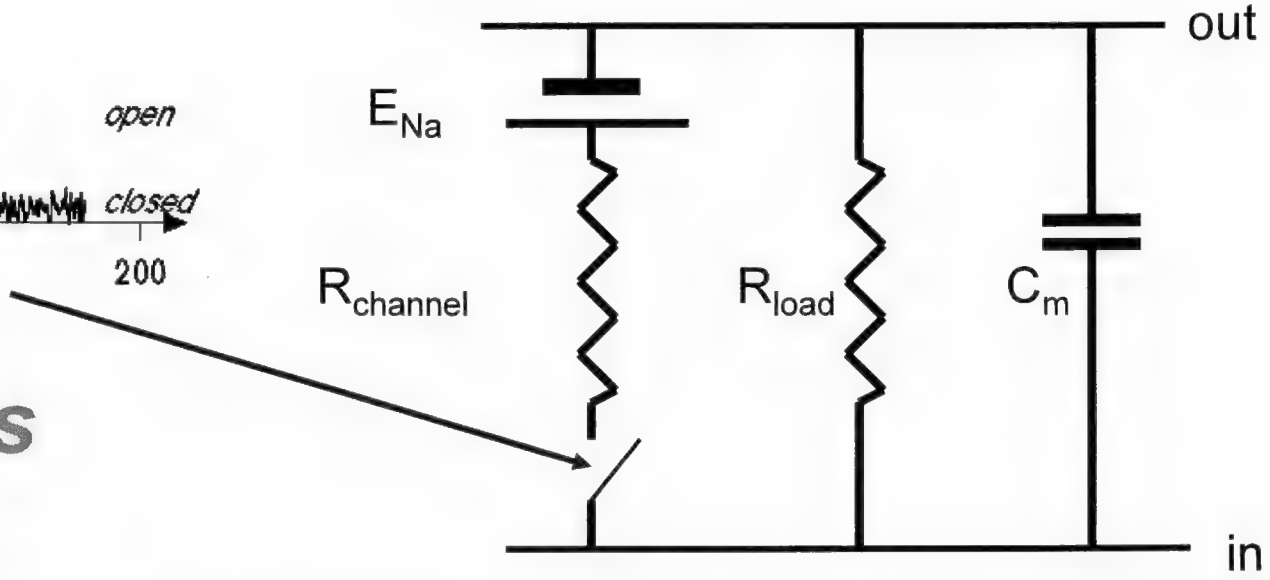
- Representational capacity is set by noise and dynamic range, and by bandwidth
- What limits these factors? Why not have faster photoreceptors, more synapses – you will see more

- Back to basics

Basic biophysics of ion channels and membranes determine signal quality



Noisy signals



$R_{channel} = 10^{11}$ Ohms

$E = 10^{-1}$ V

$R_{load} = 10^{9}$ Ohms

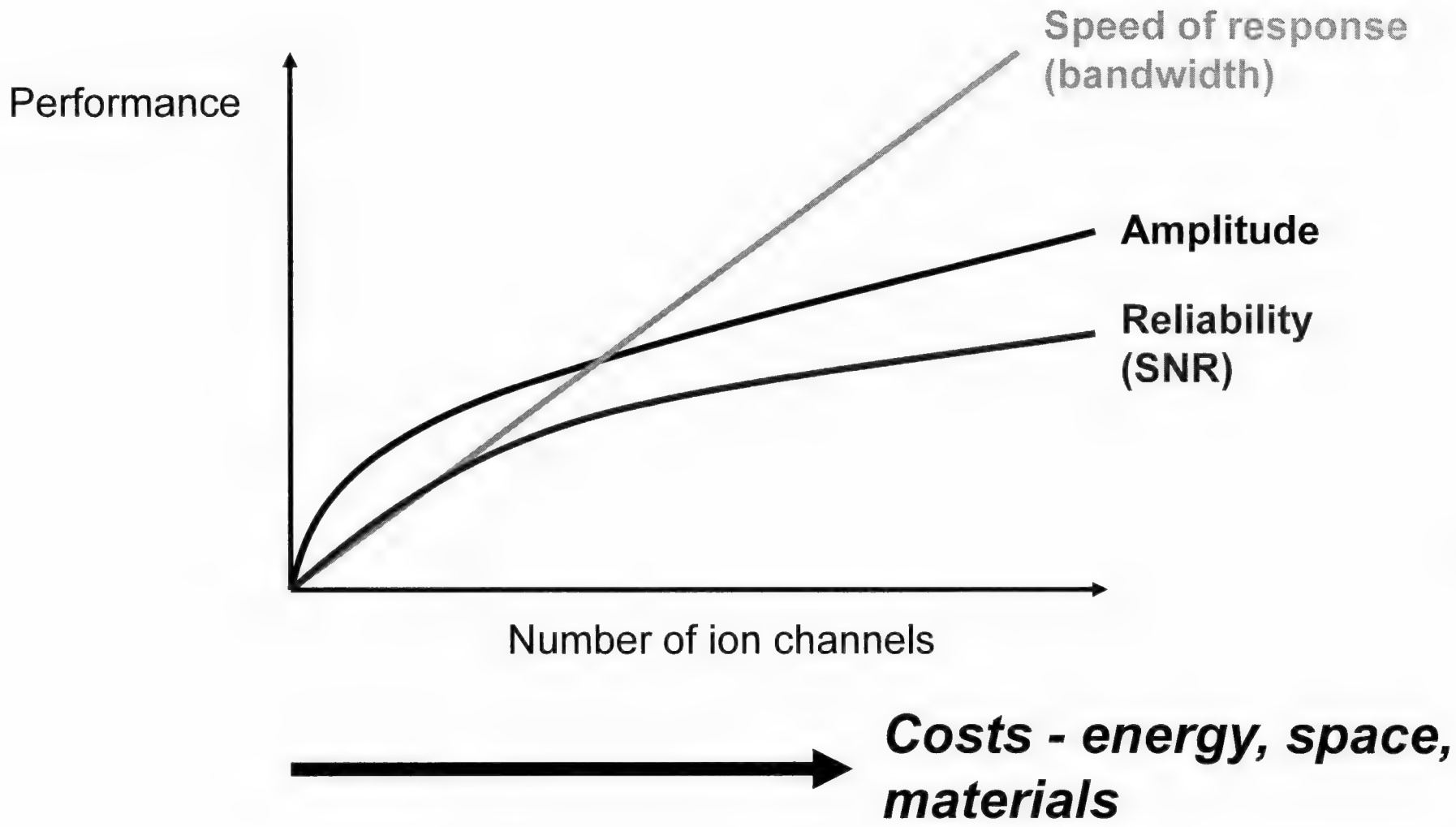
C_m is significant

R_m is high

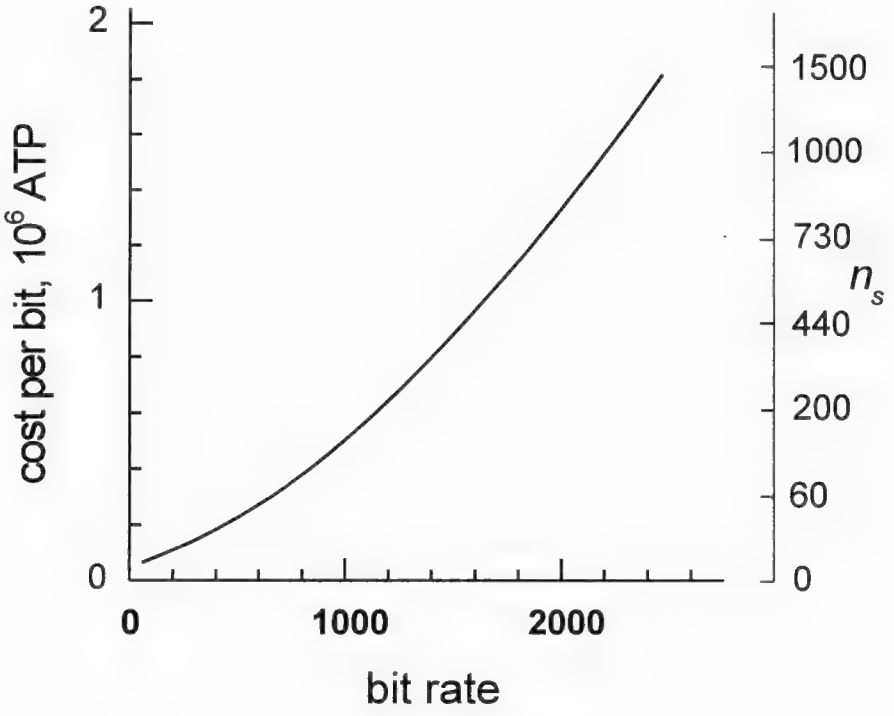
Small signals **Slow signals**

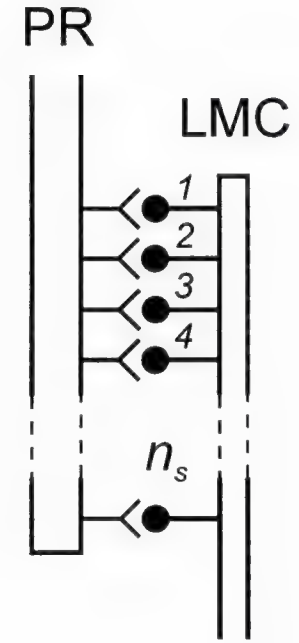
Decaying signals

Signalling performance depends on the number of ion channels generating the signal



Cost and capacity of synaptic arrays – Law of Diminishing Returns





Laughlin, Anderson de Ruyter van Steveninck, Nat Neurosci 1998

NN77 Fig. 4
Laughlin

A tale of four flies

by Jeremy Niven, John Anderson and Simon Laughlin

Sarcophaga carnaria

Calliphora vicina

Drosophila melanogaster

Drosophila virilis

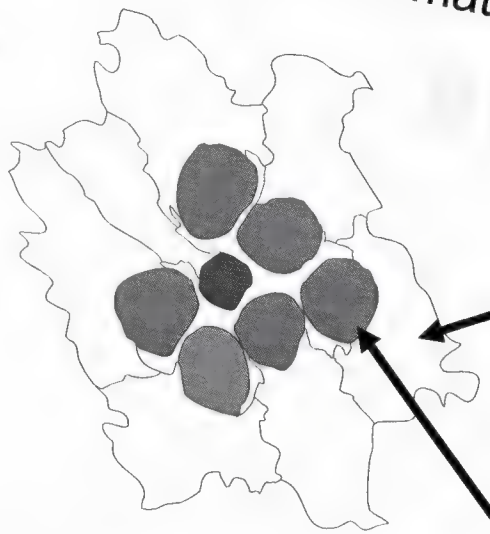
PLoS Biology 5(4) 2007

Structure of R1-6 photoreceptors

Cross-section of ommatidium

R7

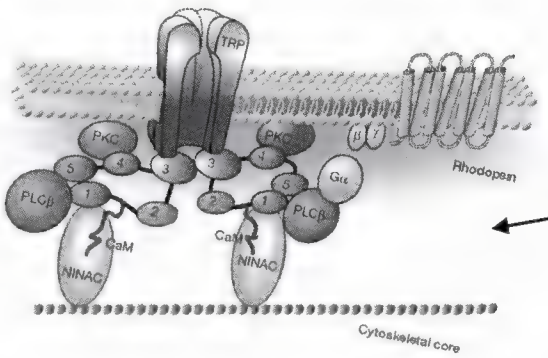
2 microns



Cell Body

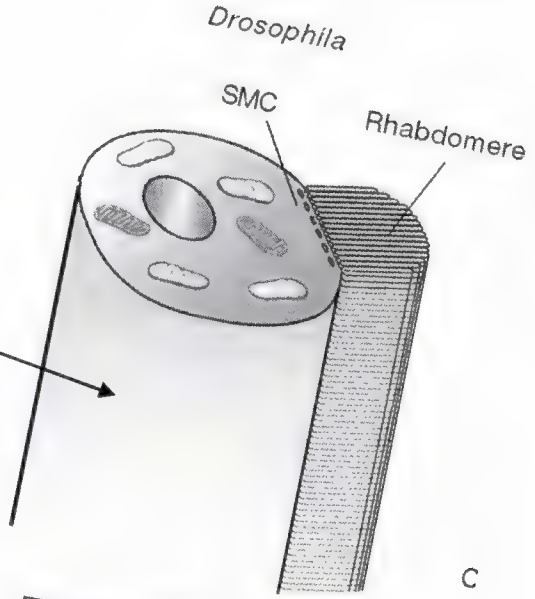
Rhabdome

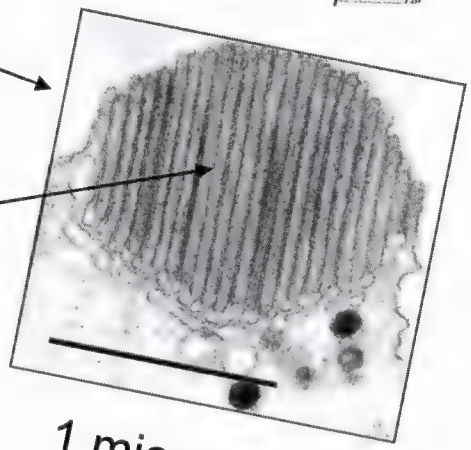
Microvillus



(Hardie & Raghu, 2001)

Drosophila

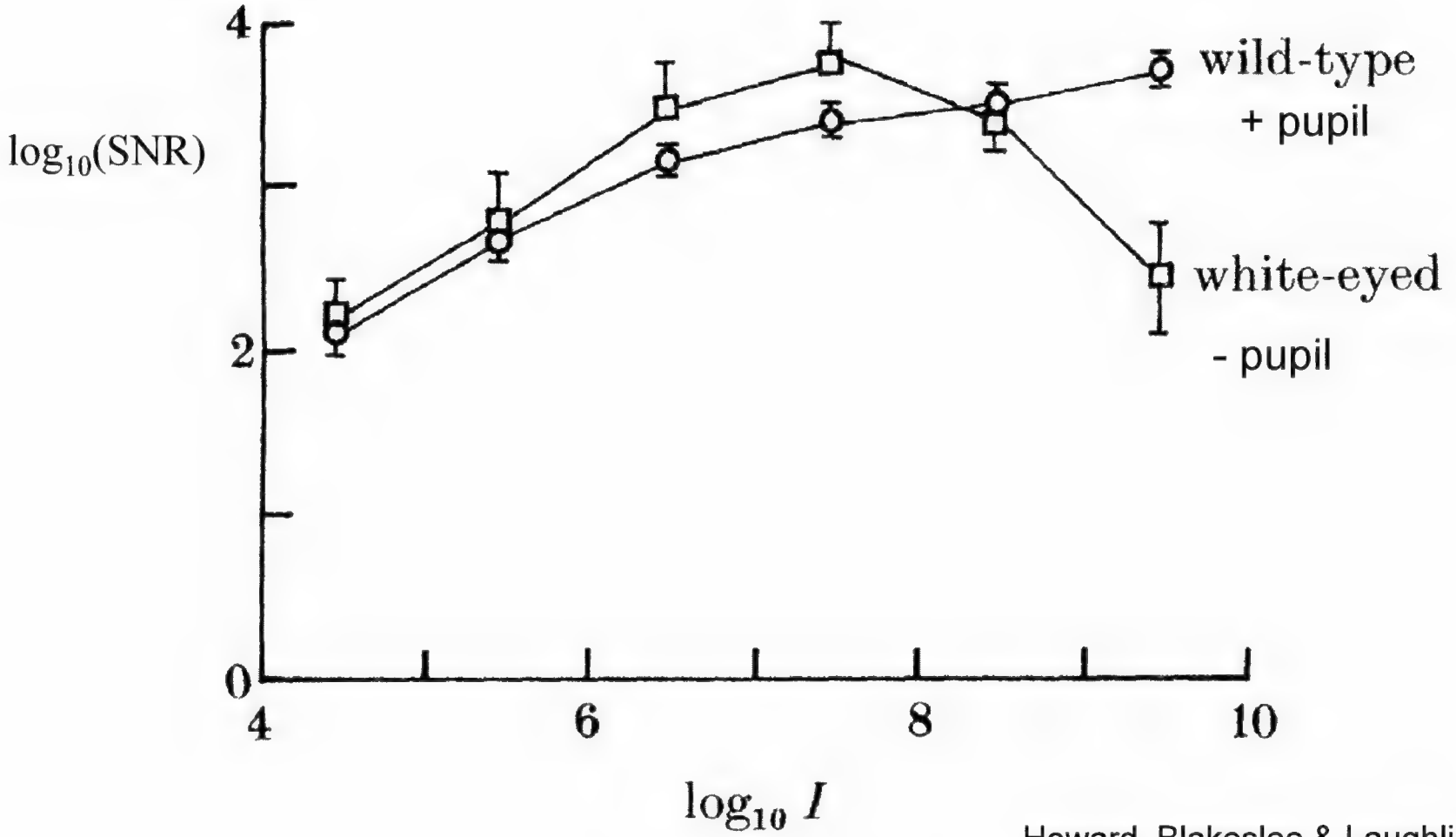




1 micron

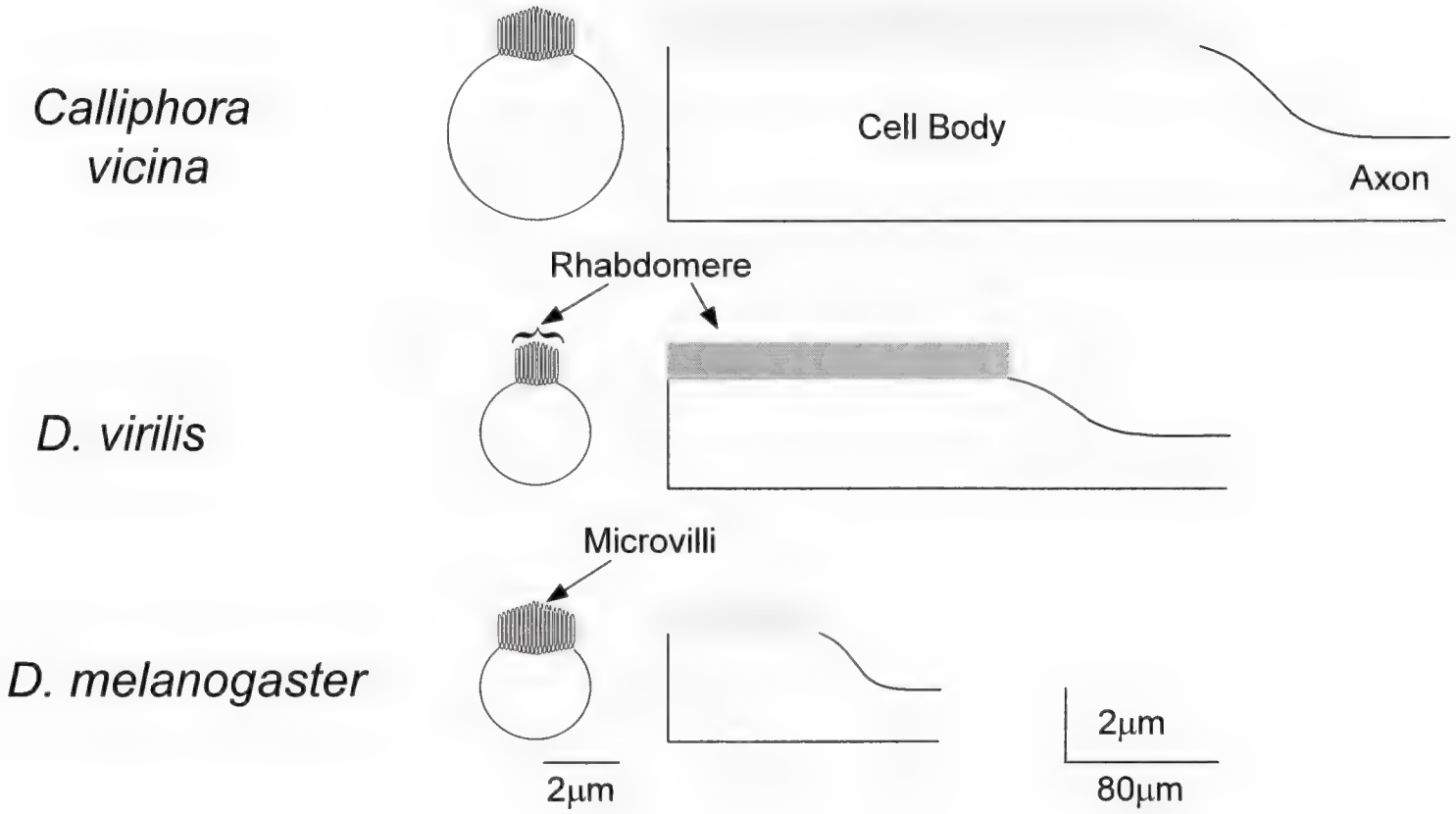
Microvillus as transduction unit

saturation depresses SNR in a pupiless mutant of *Lucilia*



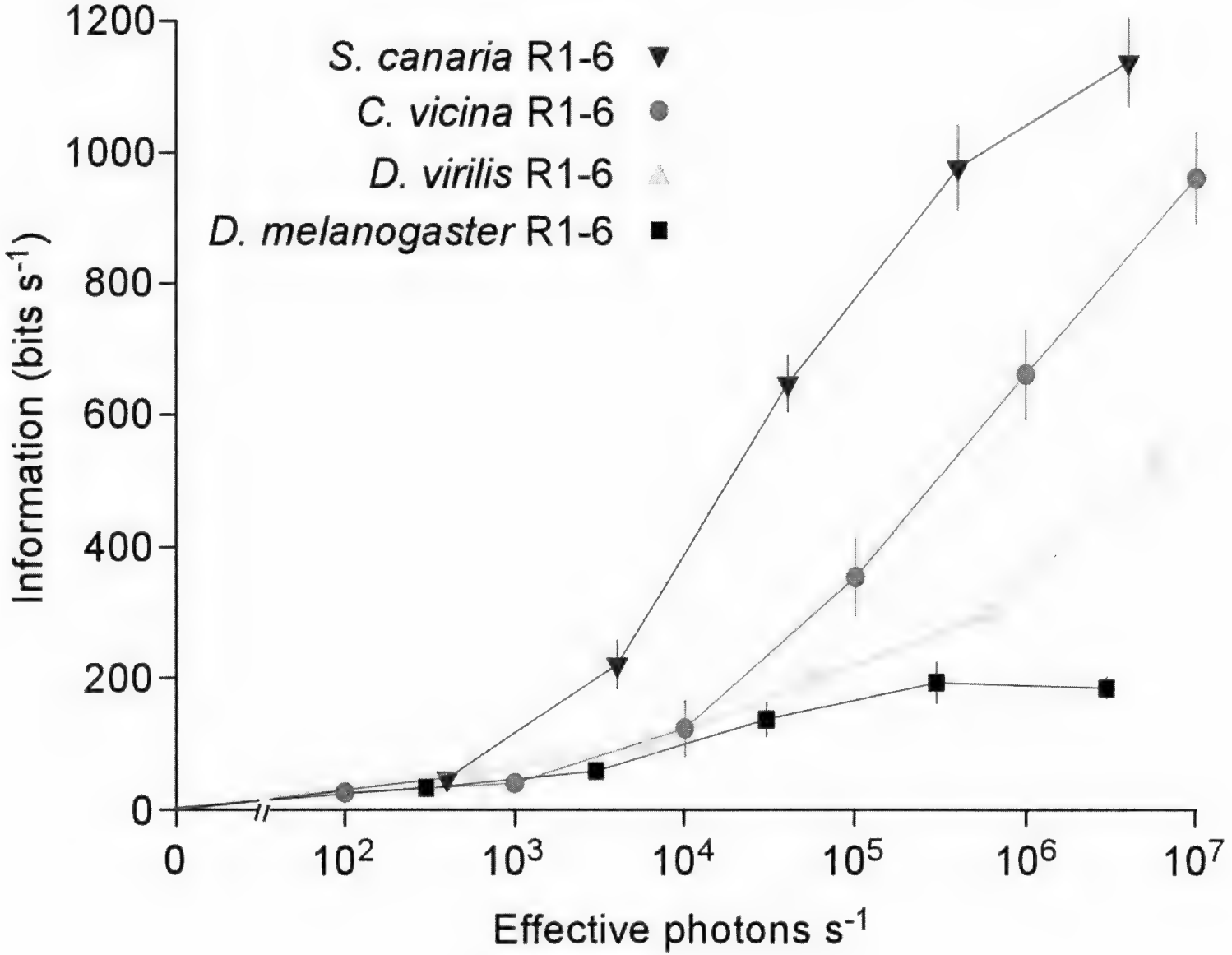
Howard, Blakeslee & Laughlin, 1987

R1-6 photoreceptors in different spp.
homologous cells – identical function
different sizes and levels of performance



Different levels of performance

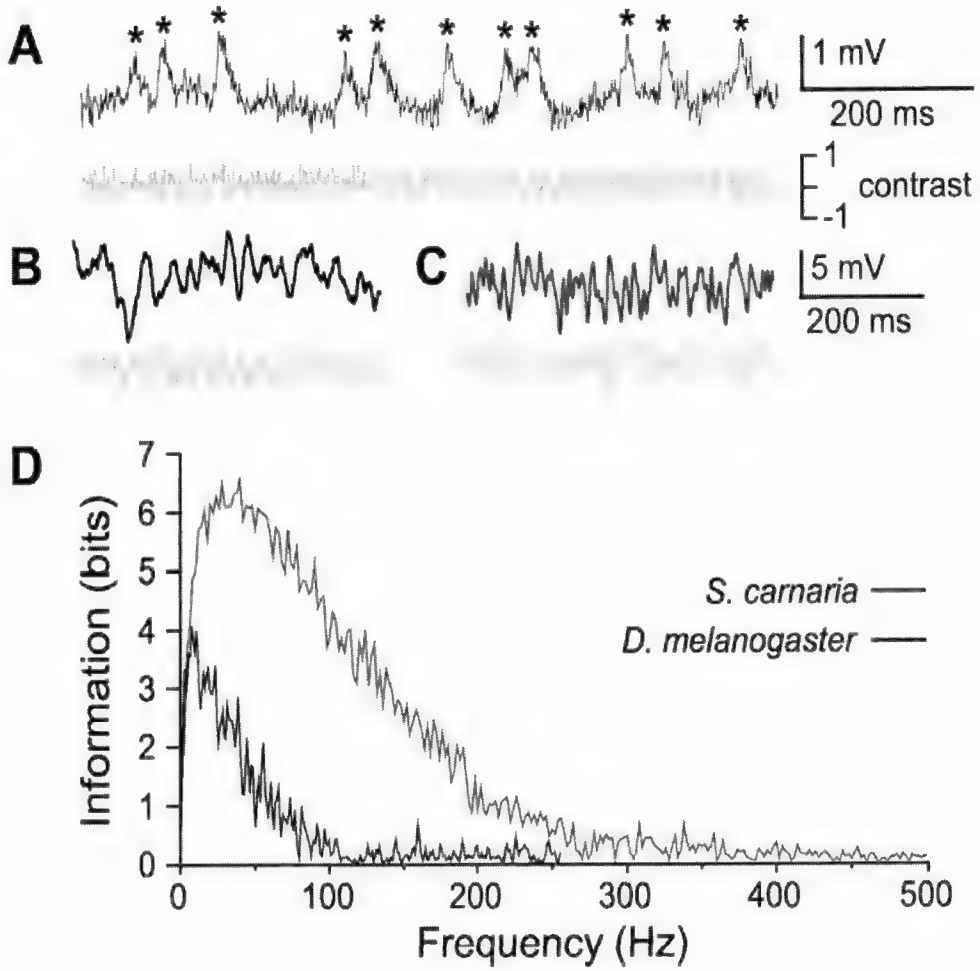
photoreceptor information rates



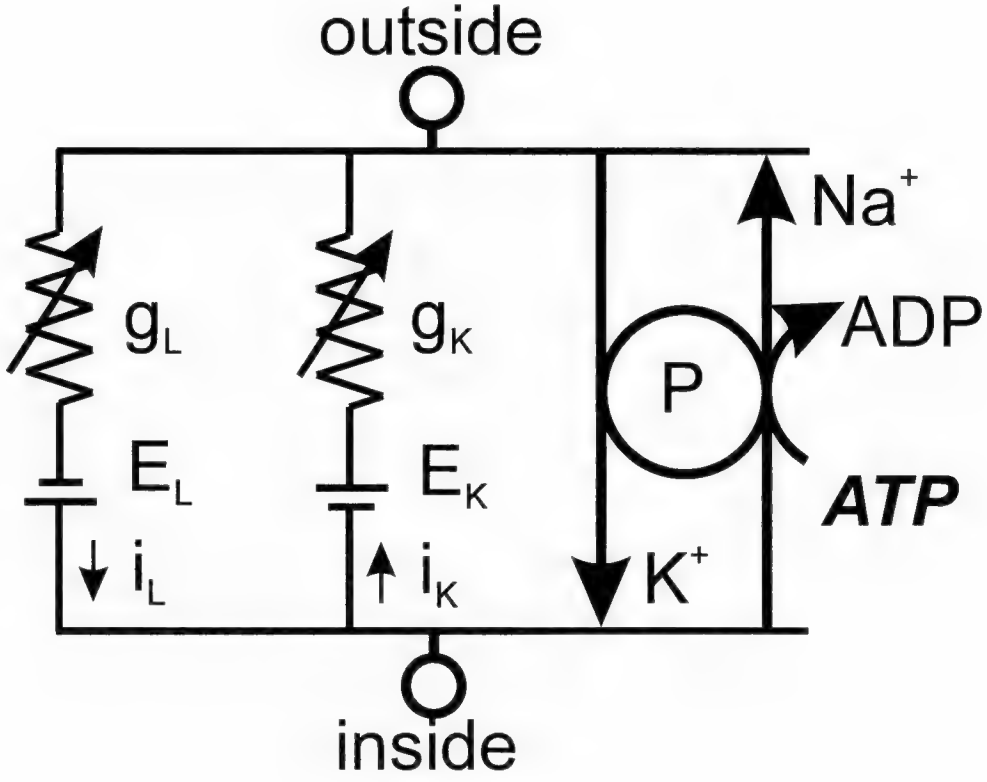
Calibrating photoreceptors by counting electrical responses single photons

Intracellular responses to "white-noise" contrast signal

Distribution of information across photoreceptor bandwidth

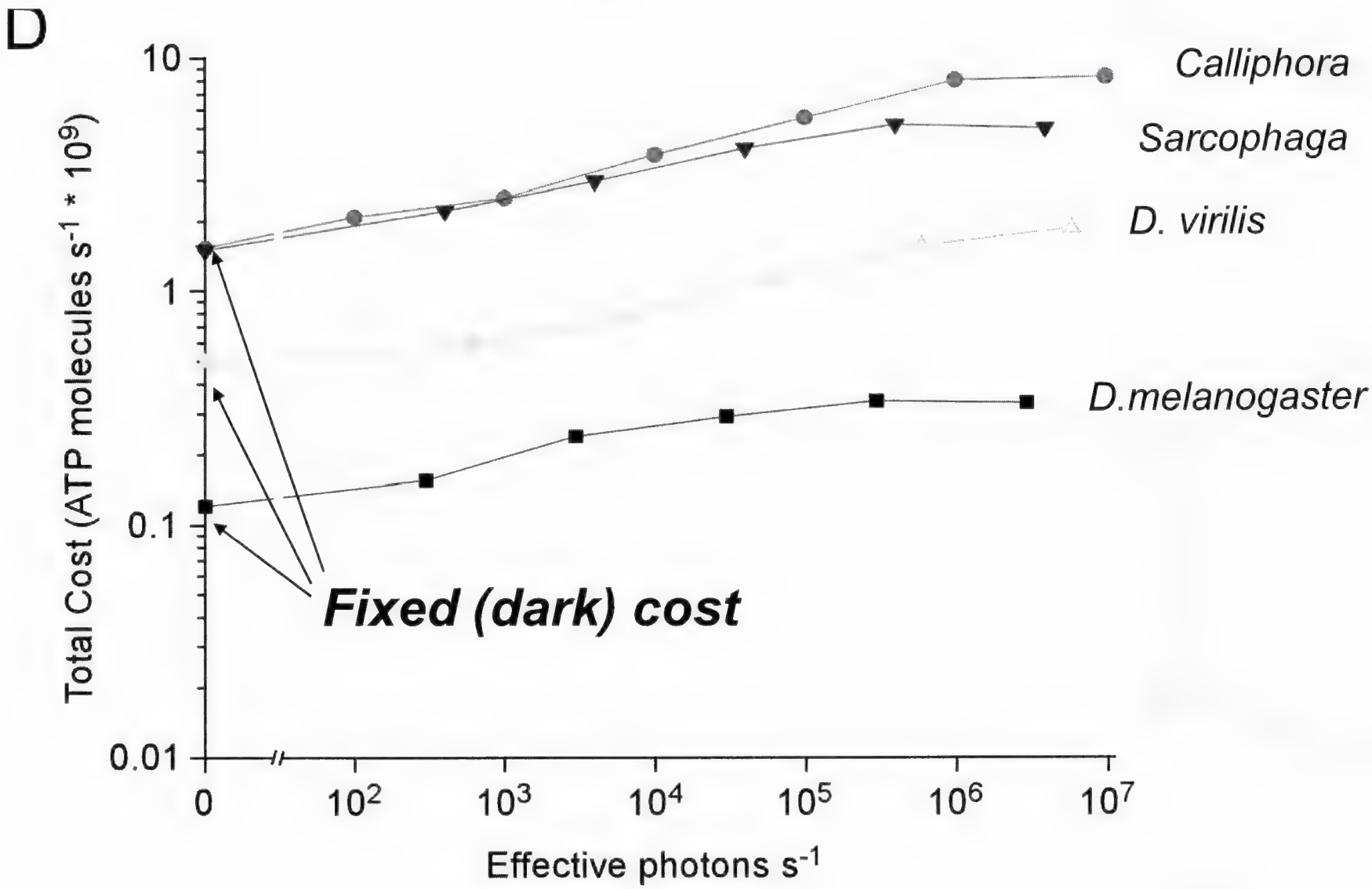


Membrane model gives energy consumption

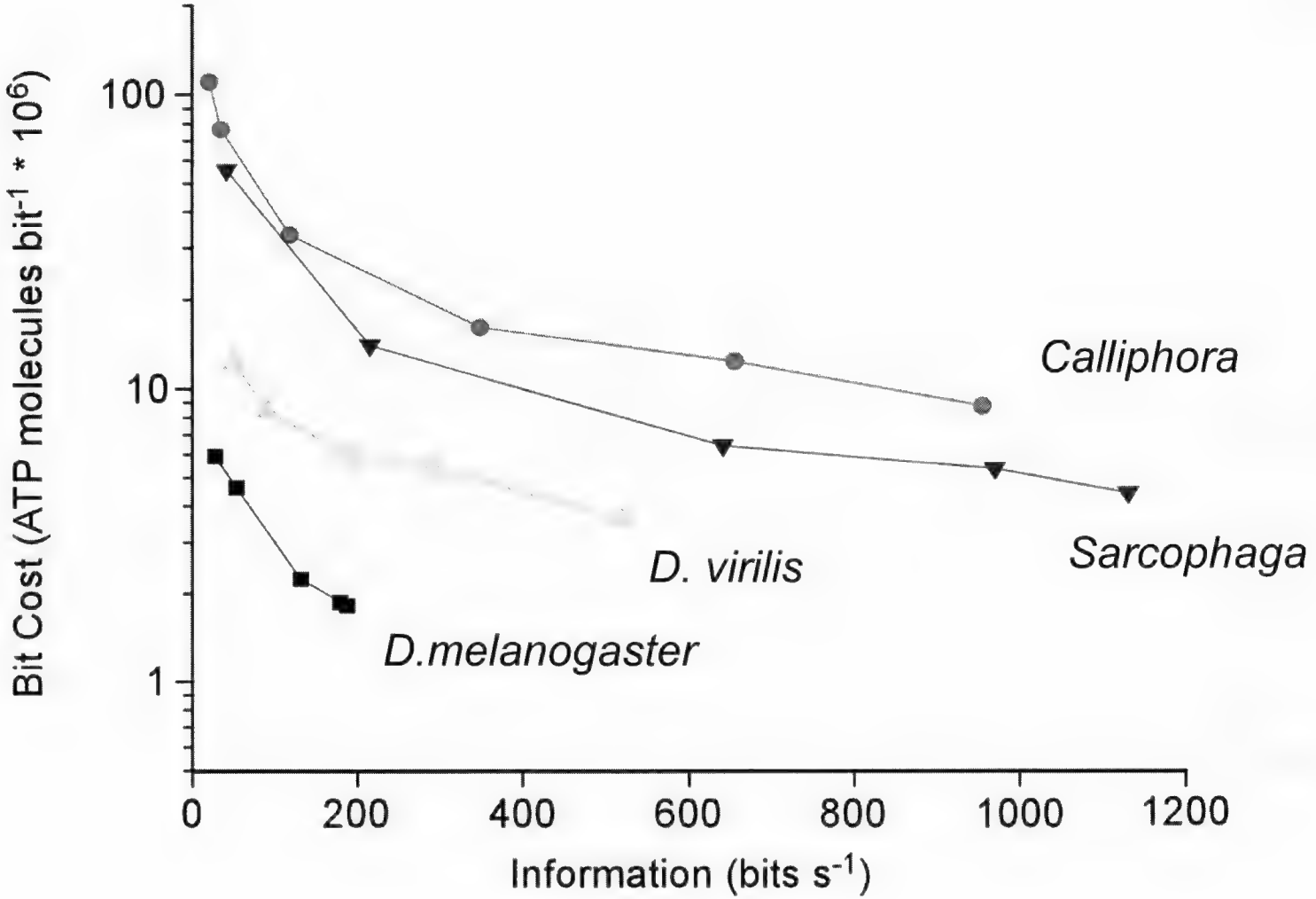


Laughlin et al, Nature Nsci 1998

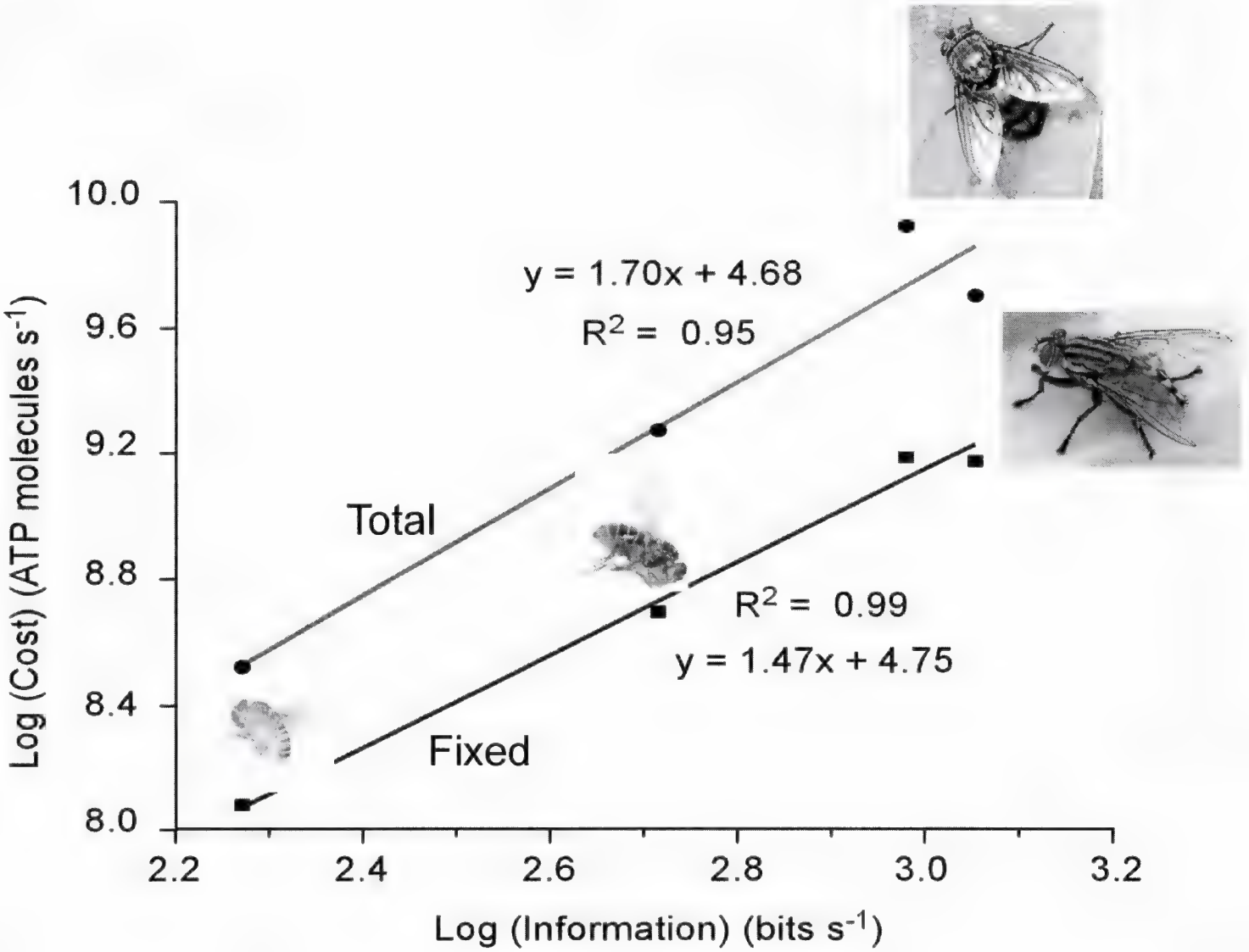
Different levels of energy consumption



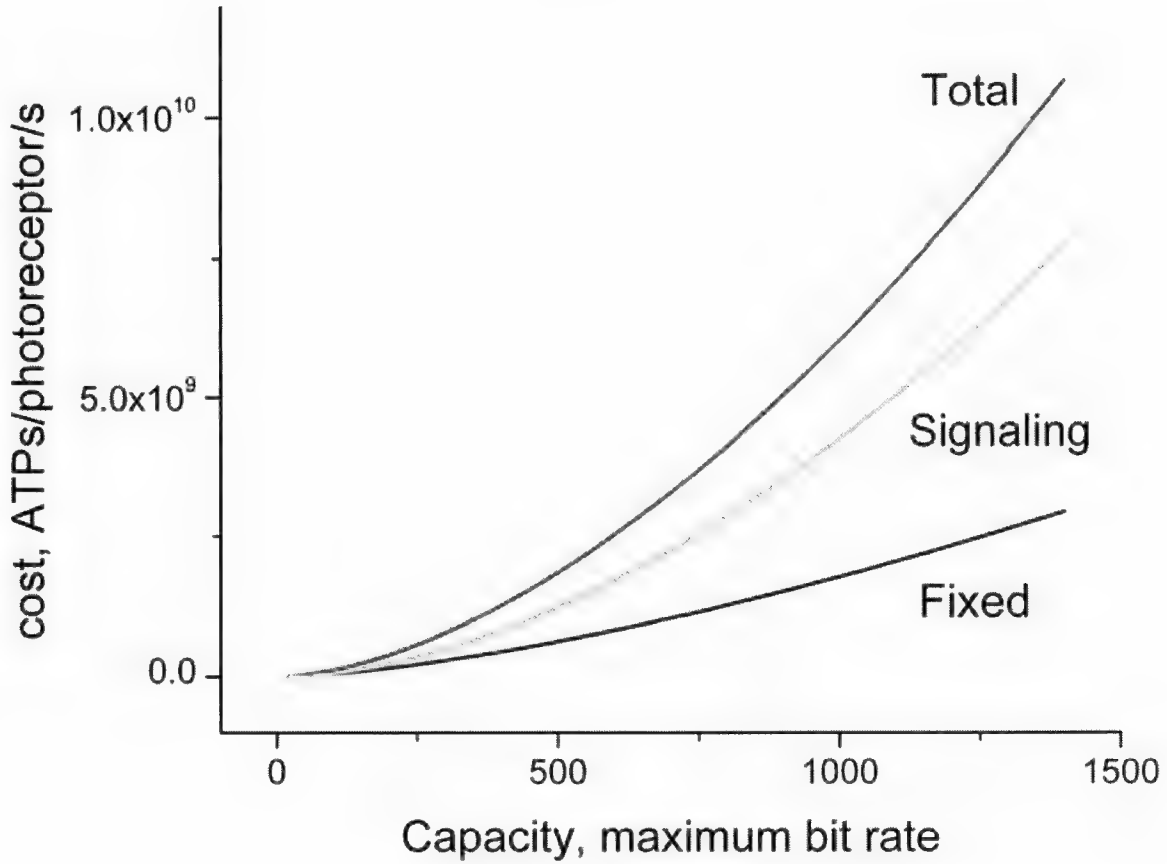
Total cost per bit
R1-6 photoreceptors of 4 *spp*

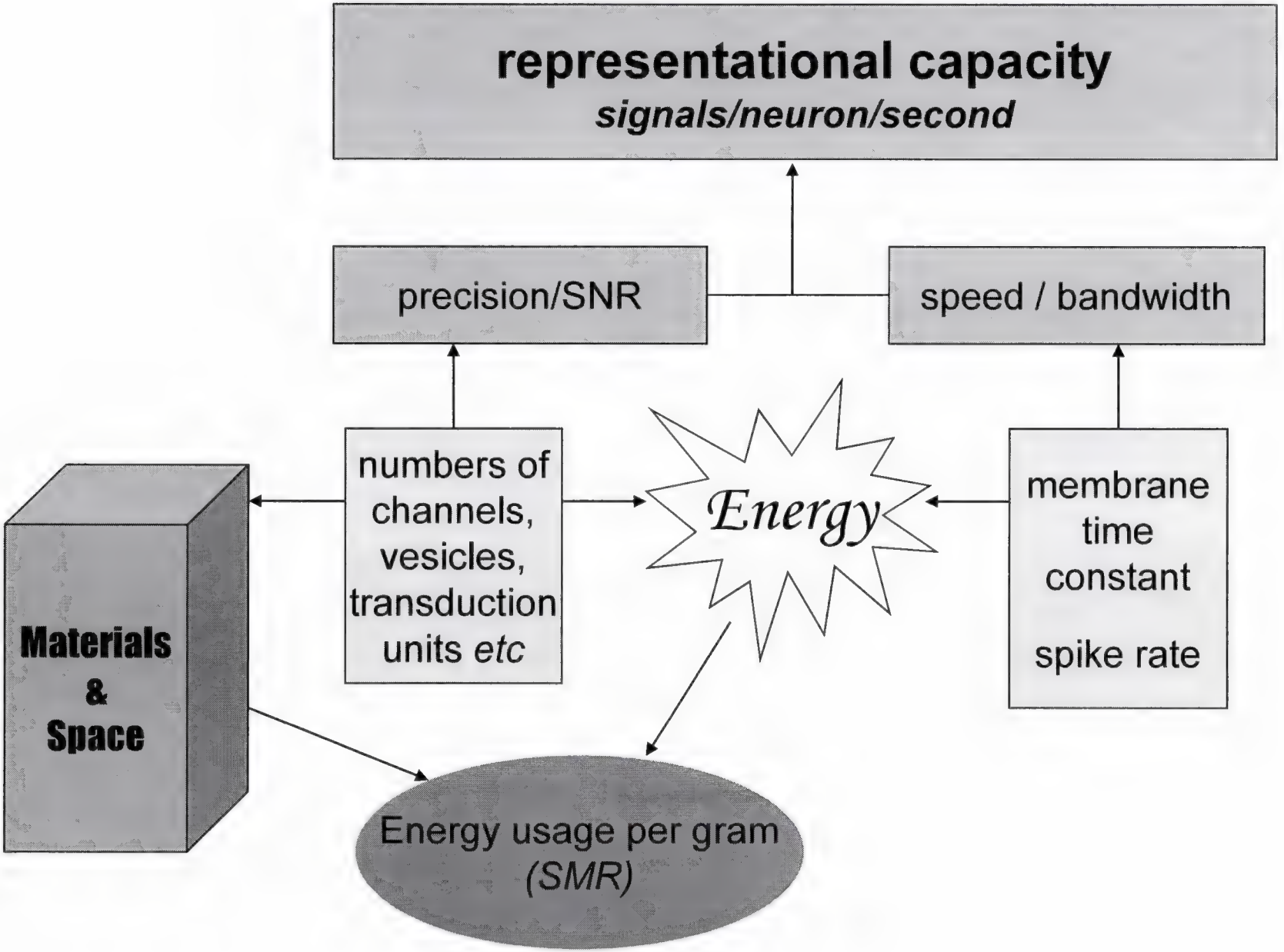


The scaling of cost with capacity



The Law of Diminishing Returns

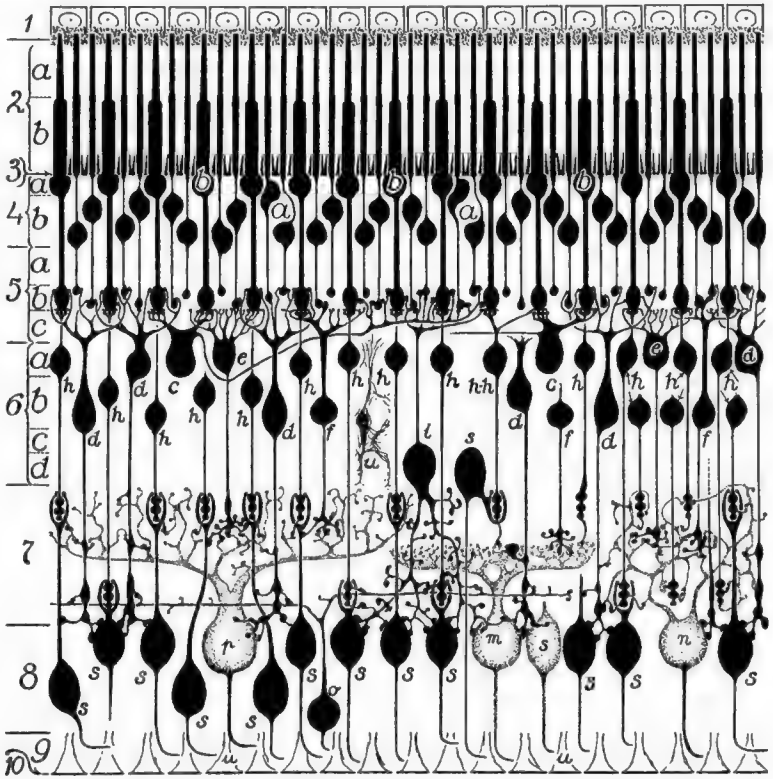



representational capacity
signals/neuron/second
precision/SNR
speed / bandwidth
numbers of channels, vesicles, transduction units etc
Energy
membrane time constant
spike rate
Materials & Space
Energy usage per gram (SMR)

Lessons from fly photoreceptors

- Neurons pay a high price for precision and speed
- Cost increases out of proportion to representational capacity
- You pay a heavy penalty for over-engineering
- Energy efficient systems will match
  - the numbers of synapses
  - the numbers of ion channels
  - the size of neurons

    to the task in hand
- And use energy efficient codes



Brisk and sluggish retinal ganglion cells in guinea pig retina

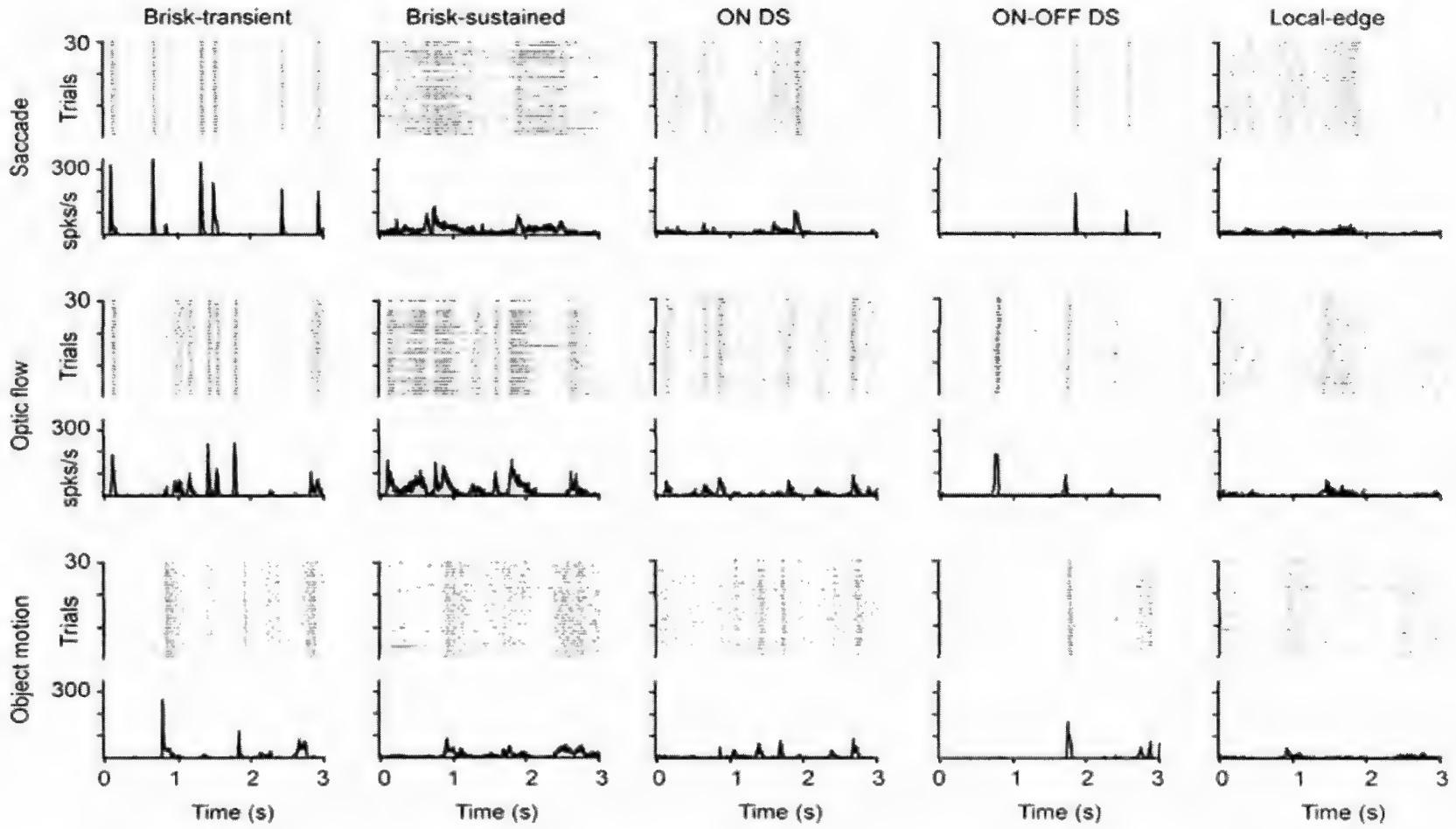
Kristin Koch, Judith McLean, Ronen Segev, Michael A. Freed, Michael J. Berry II, Vijay Balasubramanian, and Peter Sterling Curr Biol 2006

850,000 bits/s along optic nerve

64% by sluggish cells

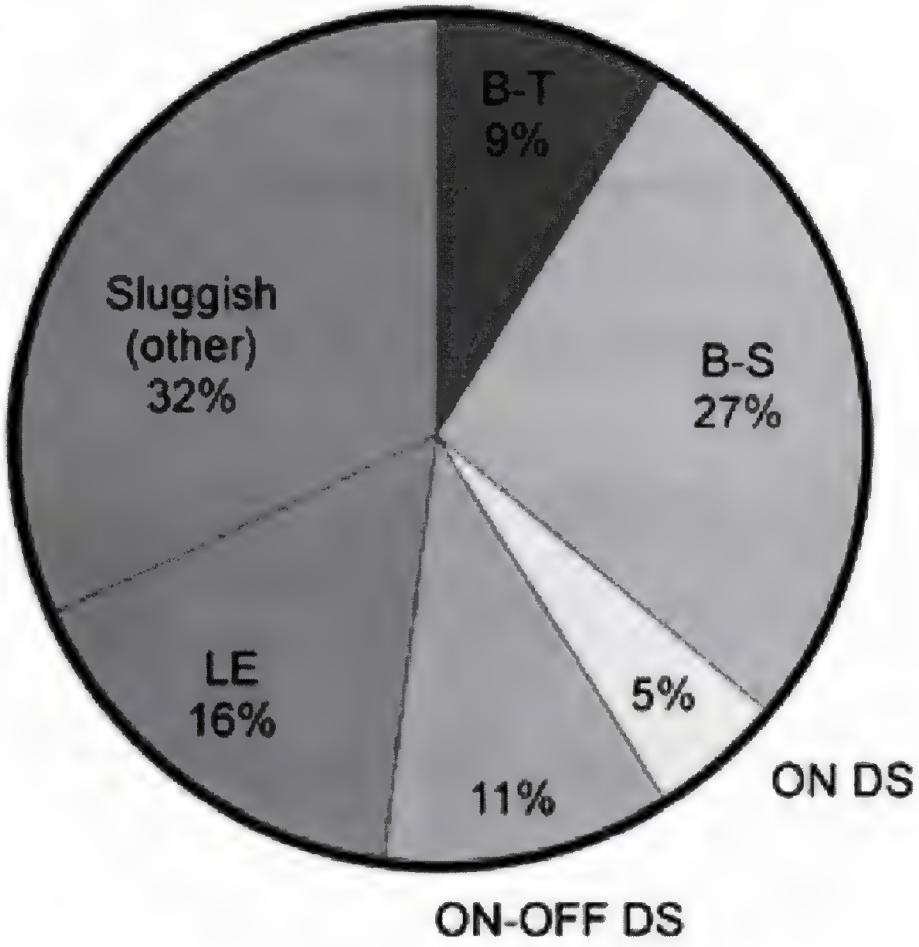
Lower performance sluggish cells operate more efficiently because they have lower fixed and signalling costs?

Spike raster plots and histograms from guinea pig retinal ganglion cells



From Koch et al 2006 see review by Vijay Balasubramanian & Peter Sterling, J Physiol 2009

% of information carried up optic nerve by different rgc classes



Vijay Balasubramanian & Peter Sterling, J Physiol 2009

Table 1. Transmitting information at low mean rates saves spikes

Comparison of three kinds of neural cables			
Type	Local edge	Brisk-transient	Super
Mean spike rate for cell	4 Hz	8 Hz	40 Hz
Bits spike^{-1}	2.1	1.8	1.1
Spikes s^{-1} for 300 bits s^{-1} traffic over cable	~140	~170	~270

Vijay Balasubramanian & Peter Sterling, J Physiol 2009

Number of bipolar cell output synapses increases with temporal precision of target ganglion cell

Sterling and Freed *Visual Neurosci* 2007

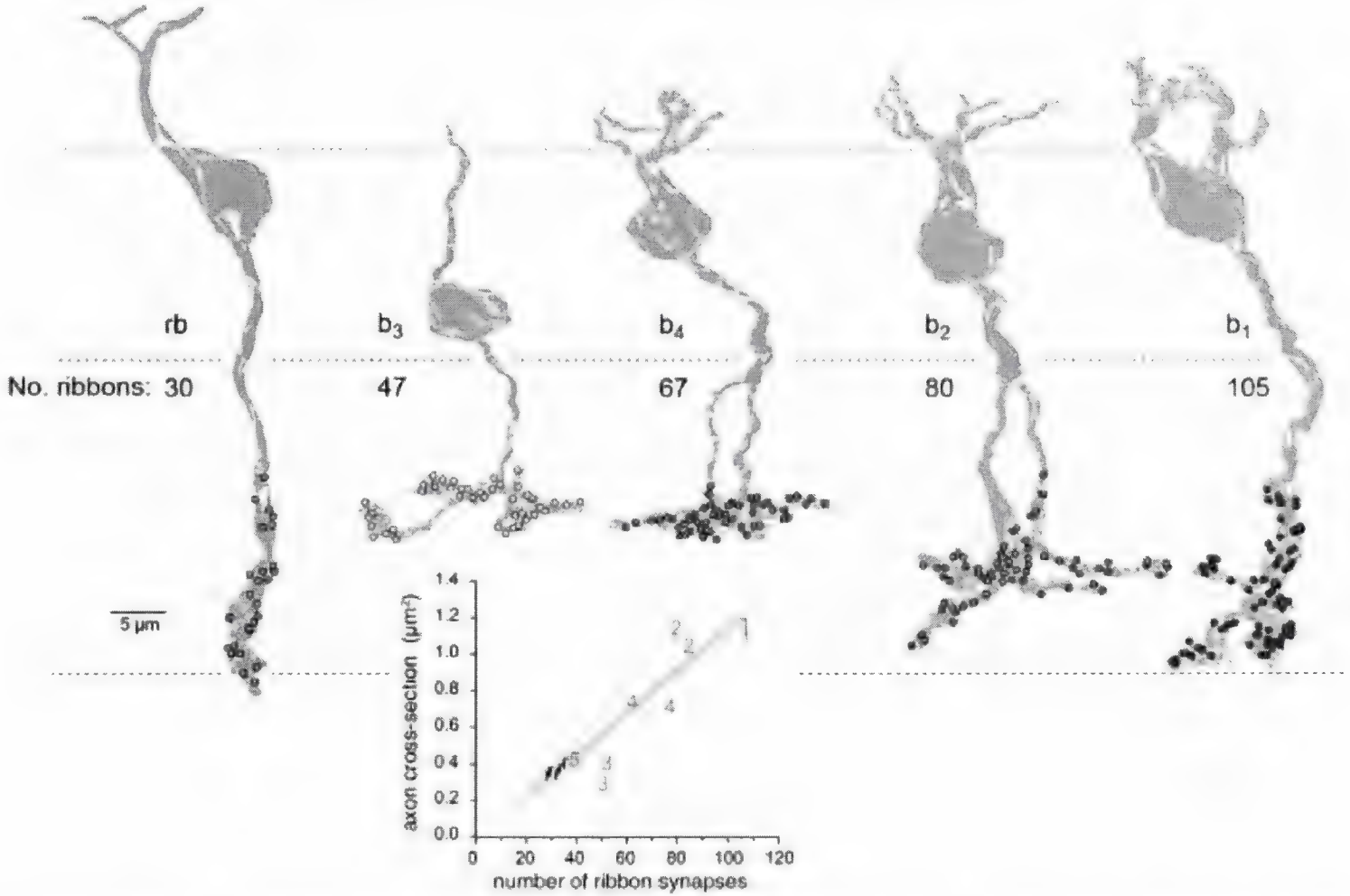
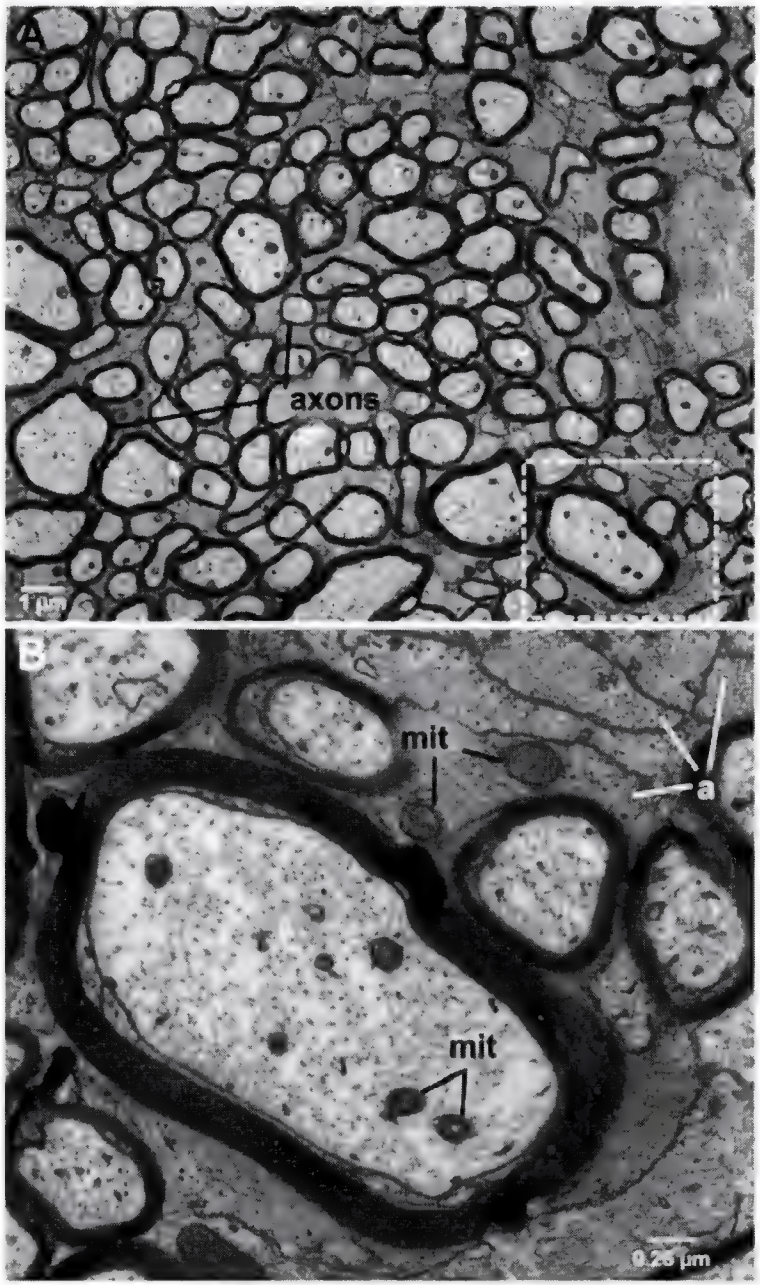
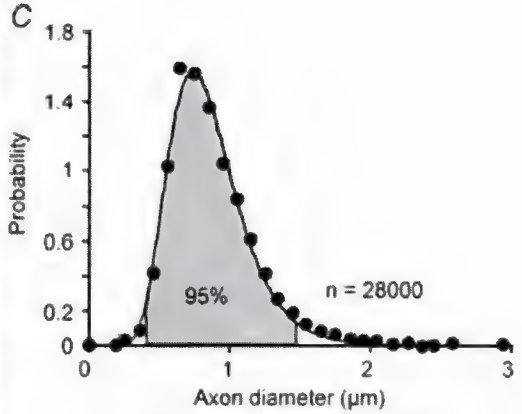


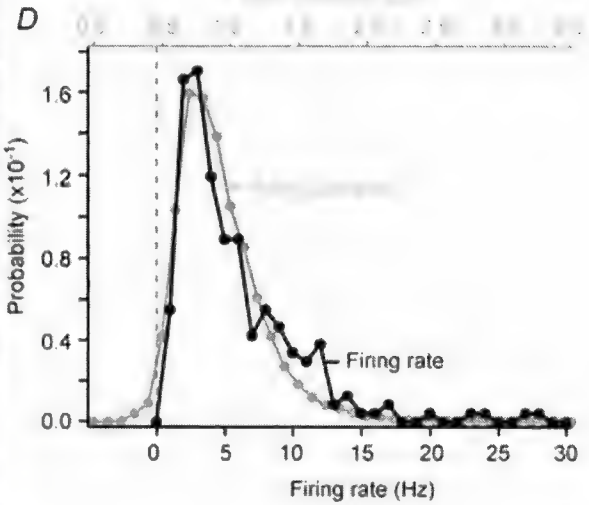
Fig. 3. Each type of ON bipolar cell expresses a different number of ribbon outputs (cat). The numbers correlate with temporal frequency, being lowest for the rod bipolar (rb) and highest for b_1, which is the transient cell that provides the main input to the brisk-transient ganglion cell. Axon cross-sectional area increases linearly with the number of ribbons. Data from Cohen and Sterling (1990).

Axon diameter *pdf* correlates with firing rate *pdf*

Guinea pig optic nerve

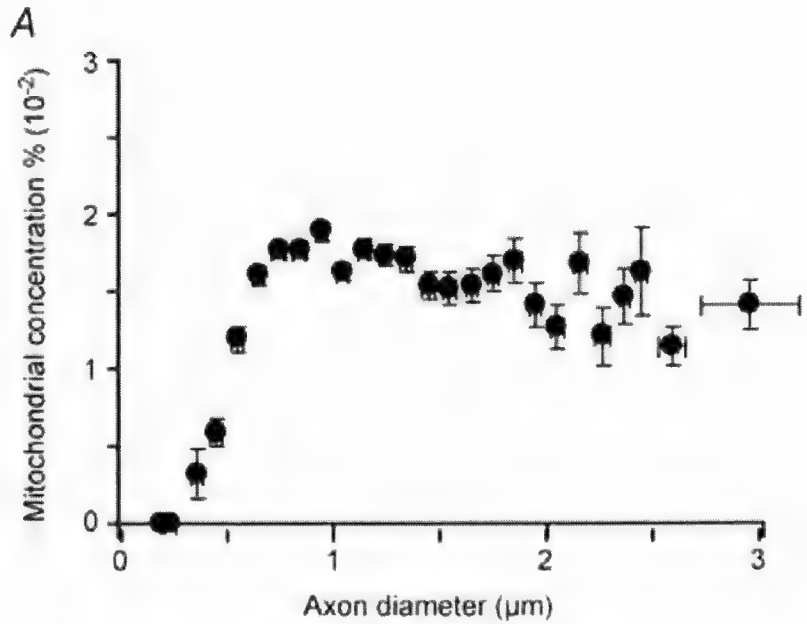






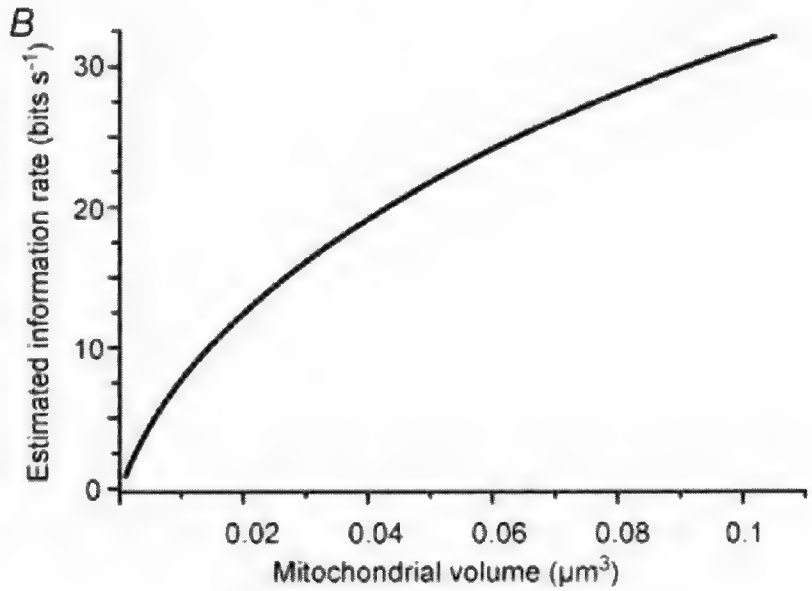
Perge et al 2009 in S & B 2009

Mitochondrial density
relatively constant across axons
of different diameter



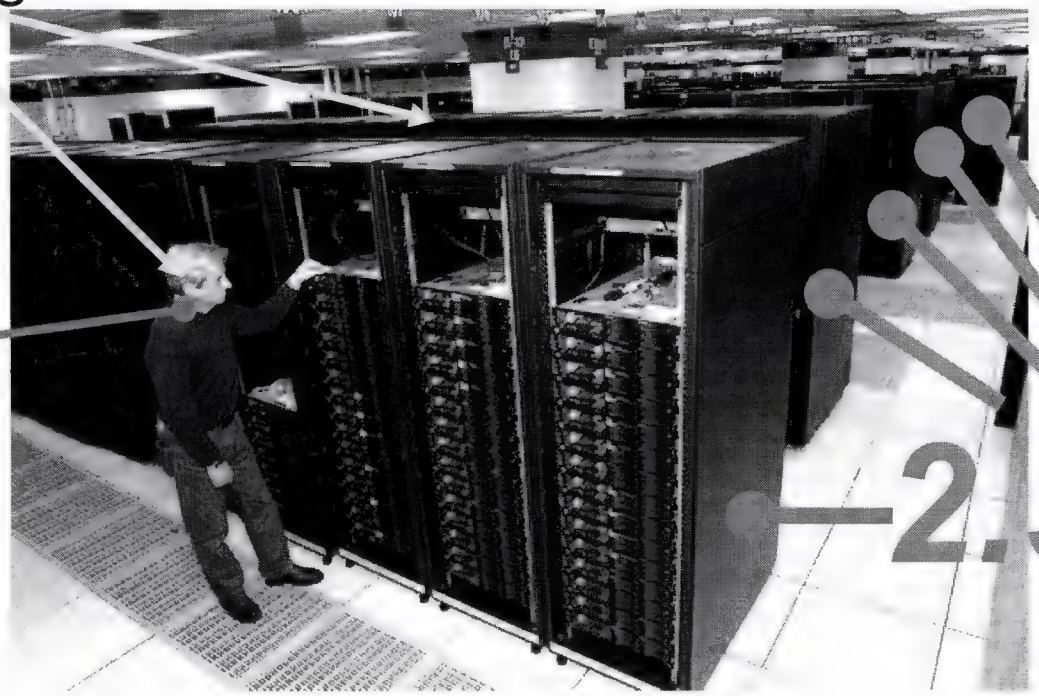
Number of mitochondria per axon
increases with bit rate –

Law of Diminishing Returns



“Joe” with IBM Road Runner Supercomputer (2007)

10^{15} events s^{-1}



15 watts

2.36 MW

Why did this cycling carry on and get places?

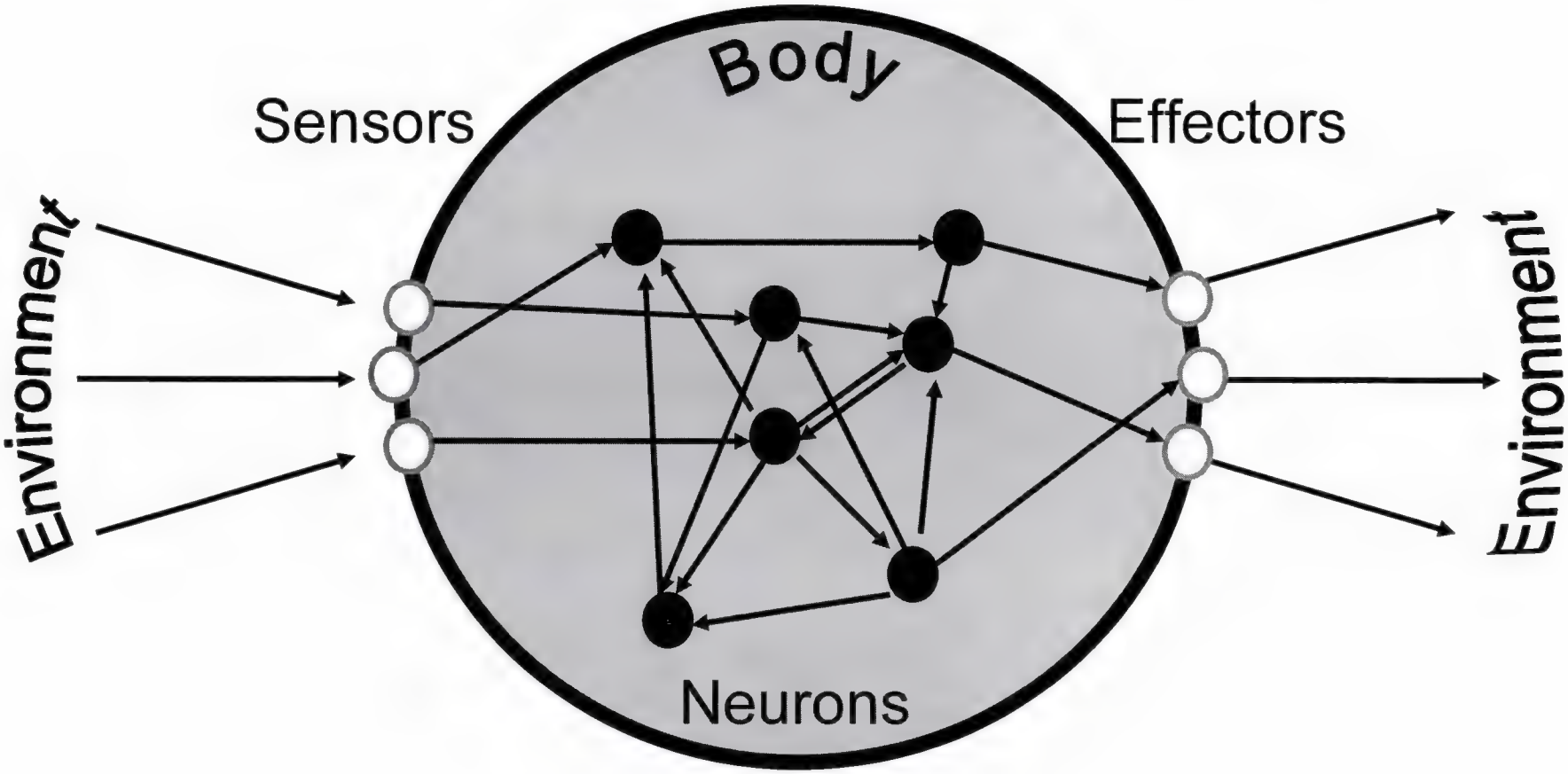
- We were OK computational neuroscientists – came up with good theories
- We were good students with great teachers - flies, guinea pigs
- Permitted definitive experiments, led us to right questions and hypotheses
- Great theory is all very well, but it has to be the appropriate theory
- Beware of futile cycles

And we are still going to college

- Coding really natural signals
- Is retinal adaptation optimising energy efficiency?
- Energy efficient spiking neurons and circuits
- Analogue vs digital
- Synaptic noise and computation

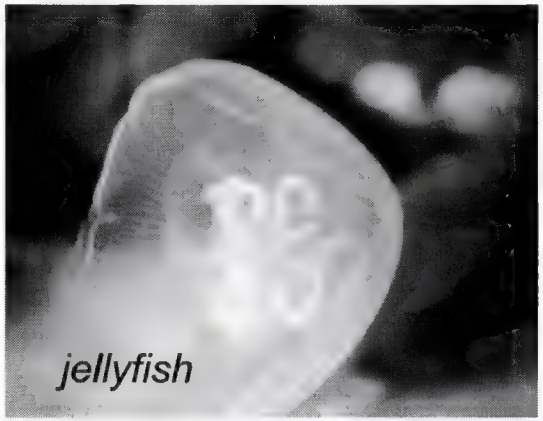
$$C(R,dt) = (-Rdt \log2 (Rdt) - (1 - Rdt) \log2 (1 - Rdt))/dt \text{ bits s}-1$$

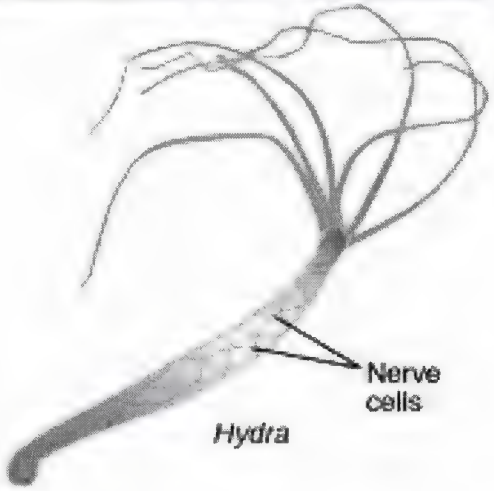
Theoretical connectomics: uncovering principles of brain design



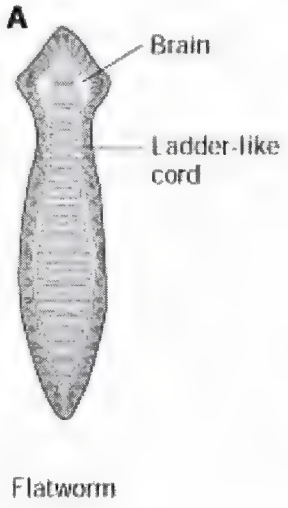
What is the appropriate level of abstraction?

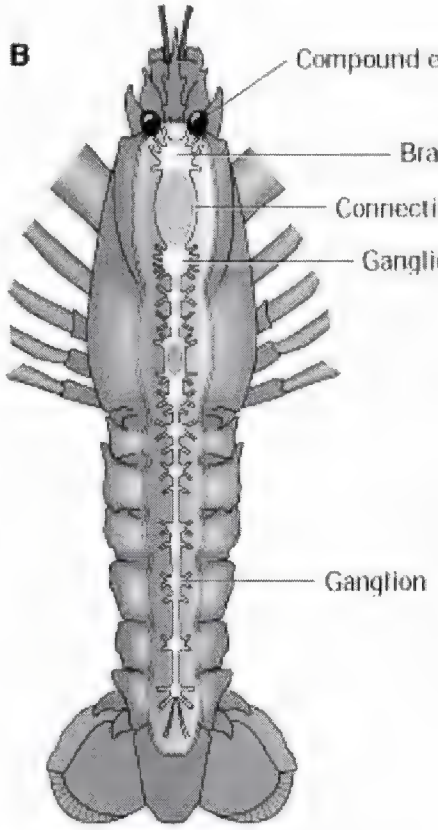
Distributed vs. centralized nervous systems (Cajal 1899)

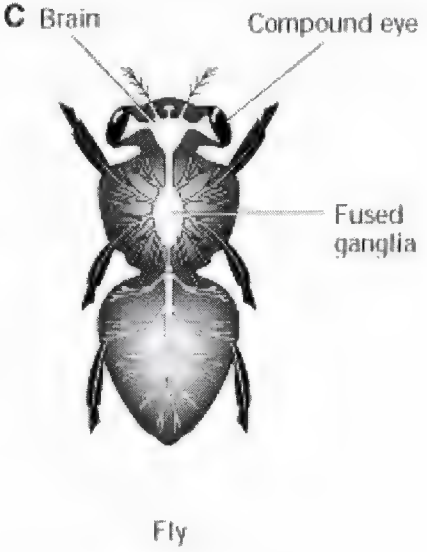




10^3 neurons



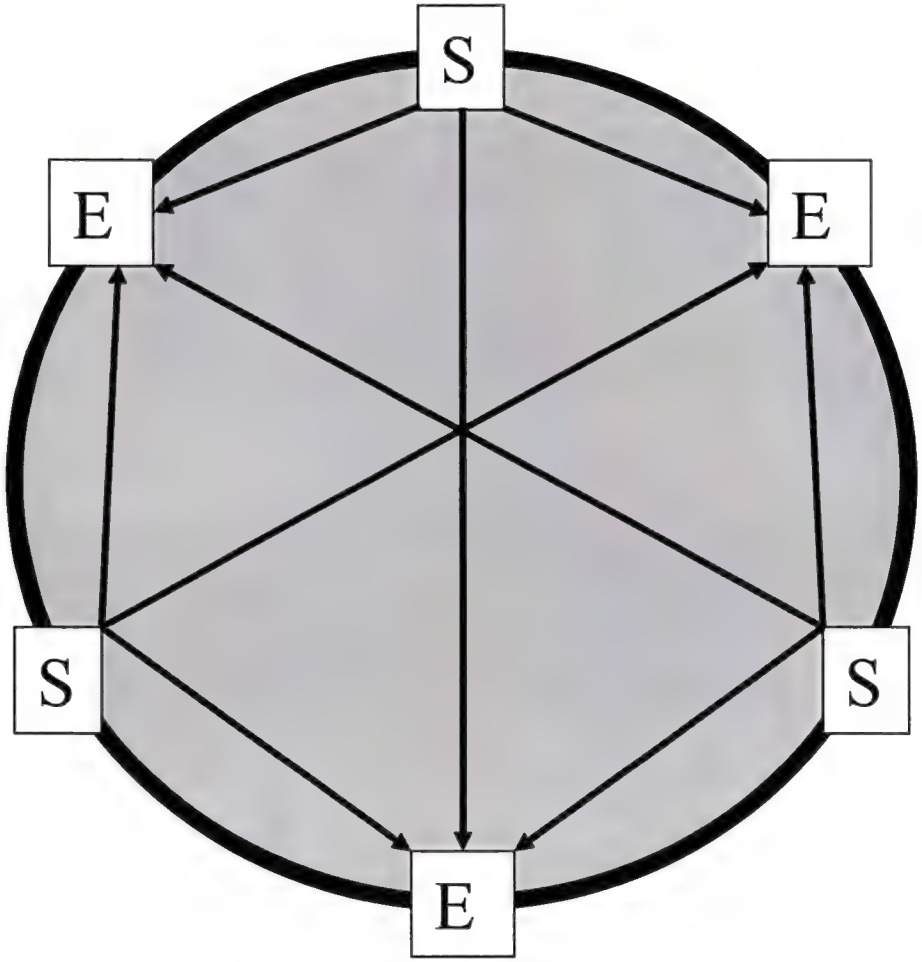




10^5 neurons

Which design is less costly?

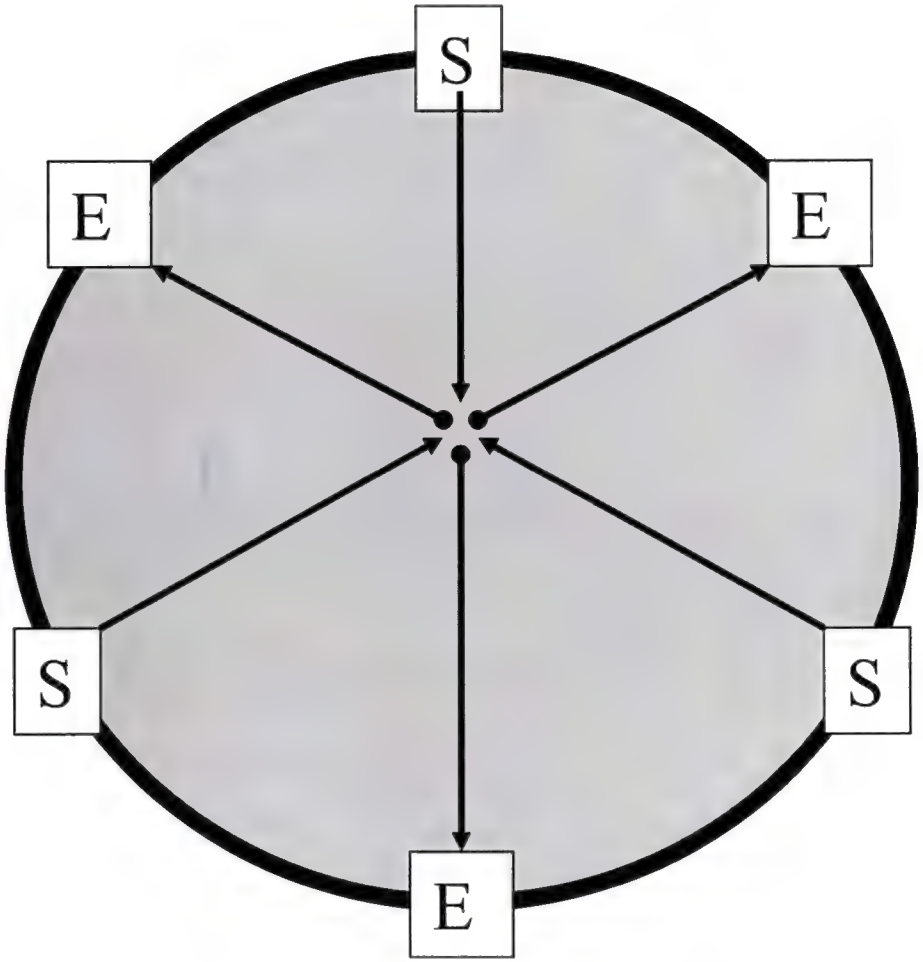
Distributed nervous system



$N_{Wires} = N_S N_E$

$N_S = N_E = 10^6$: 10^{12}

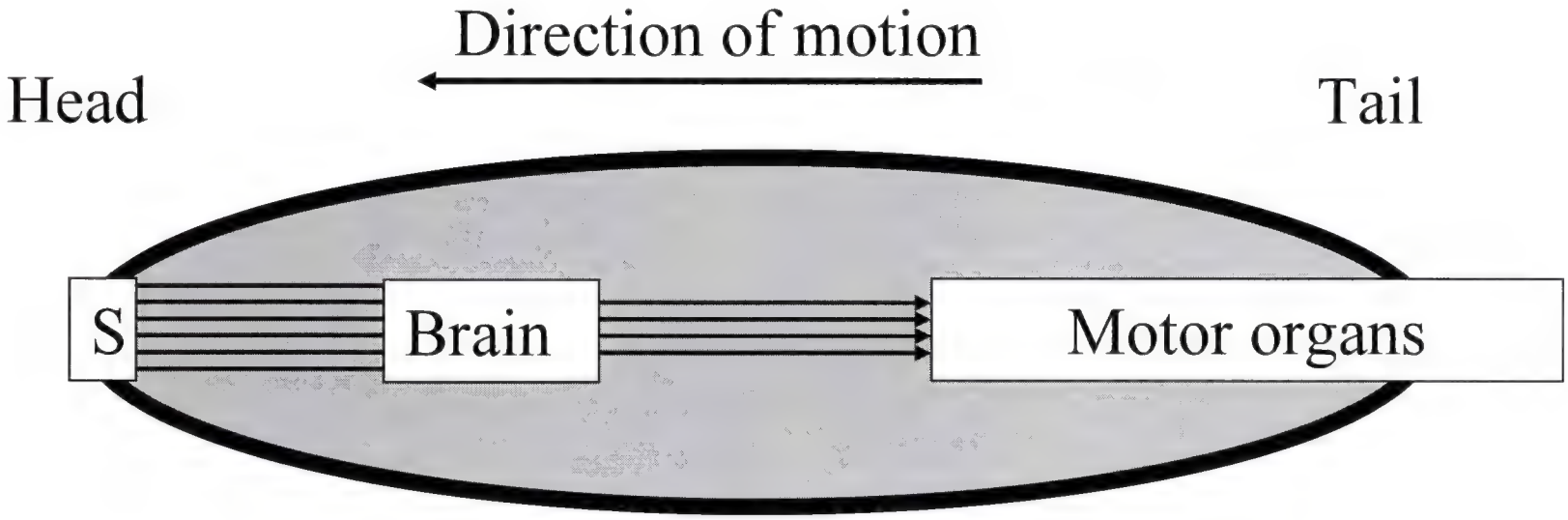
Centralized nervous system



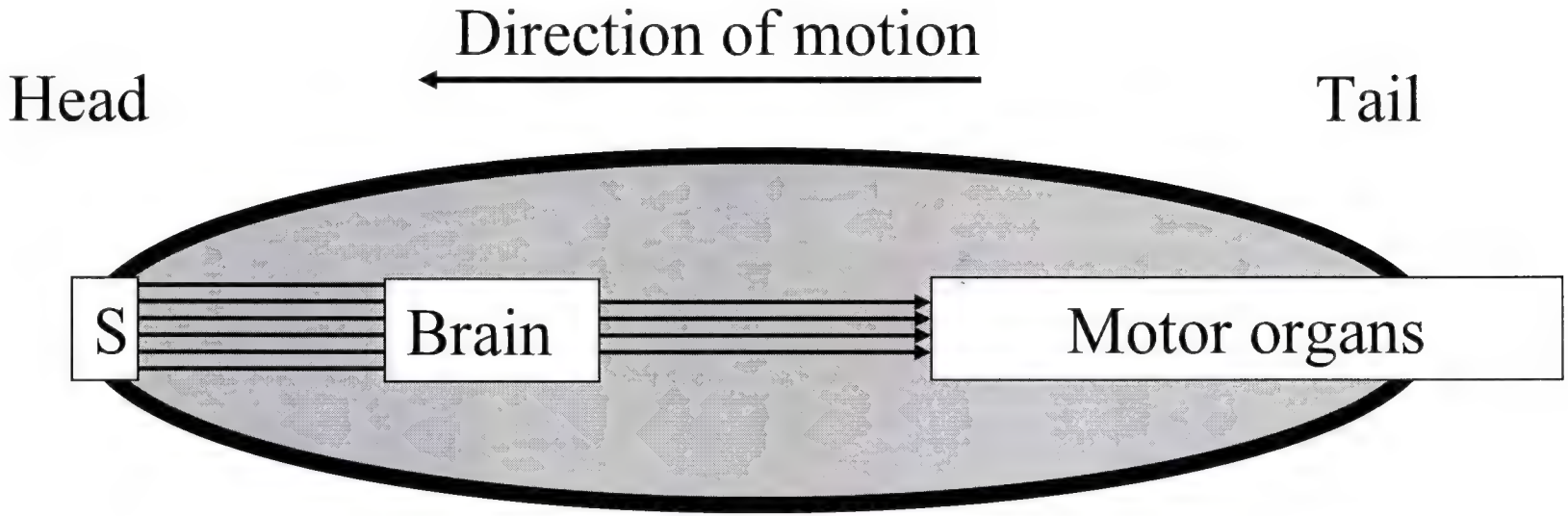
$N_{Wires} = N_S + N_E$

2×10^6

What determines brain placement?



What determines brain placement?

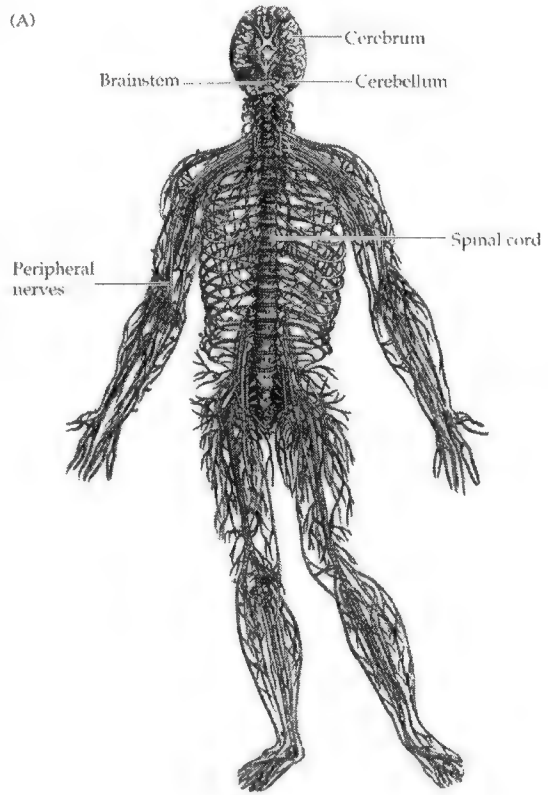


Brain location is biased towards the dominant source of connections

Numbers of connections to the human brain

Anterior: Cranial nerves

Olfactory	10,000K
Optic	2,000K
Oculomotor	60K
Trochlear	6K
Trigeminal	300K
Abducens	14K
Facial	20K
Cochlear	60K
Vestibular	40K
Glossopharyngeal	7K
Vagus	70K
Accessory	7K
Hypoglossal	15K



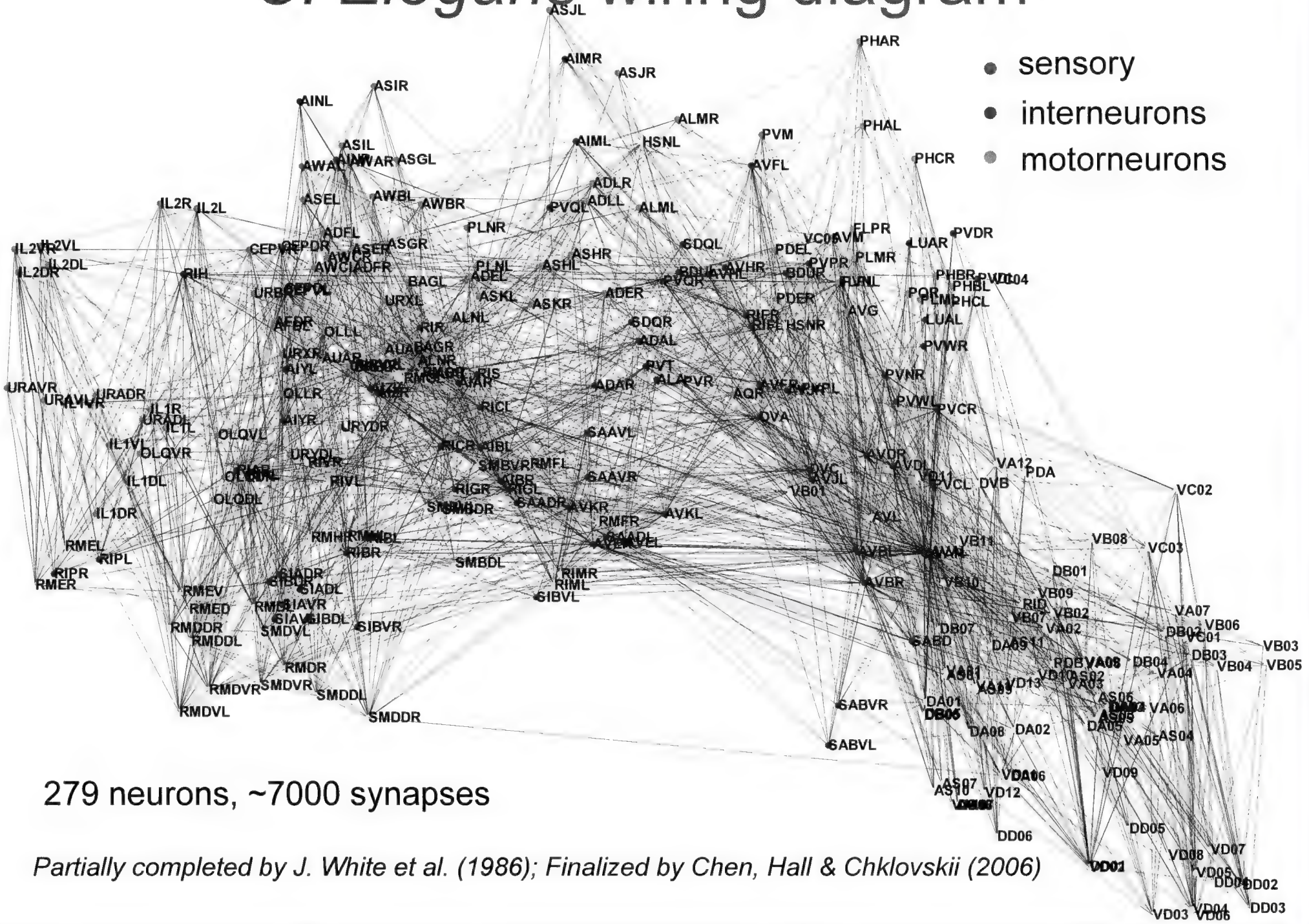
Posterior: Spinal cord

Dorsal	2,000K
Ventral	400K

Anterior/posterior ratio > 1 is consistent with forward brain placement

Cherniak, 1994

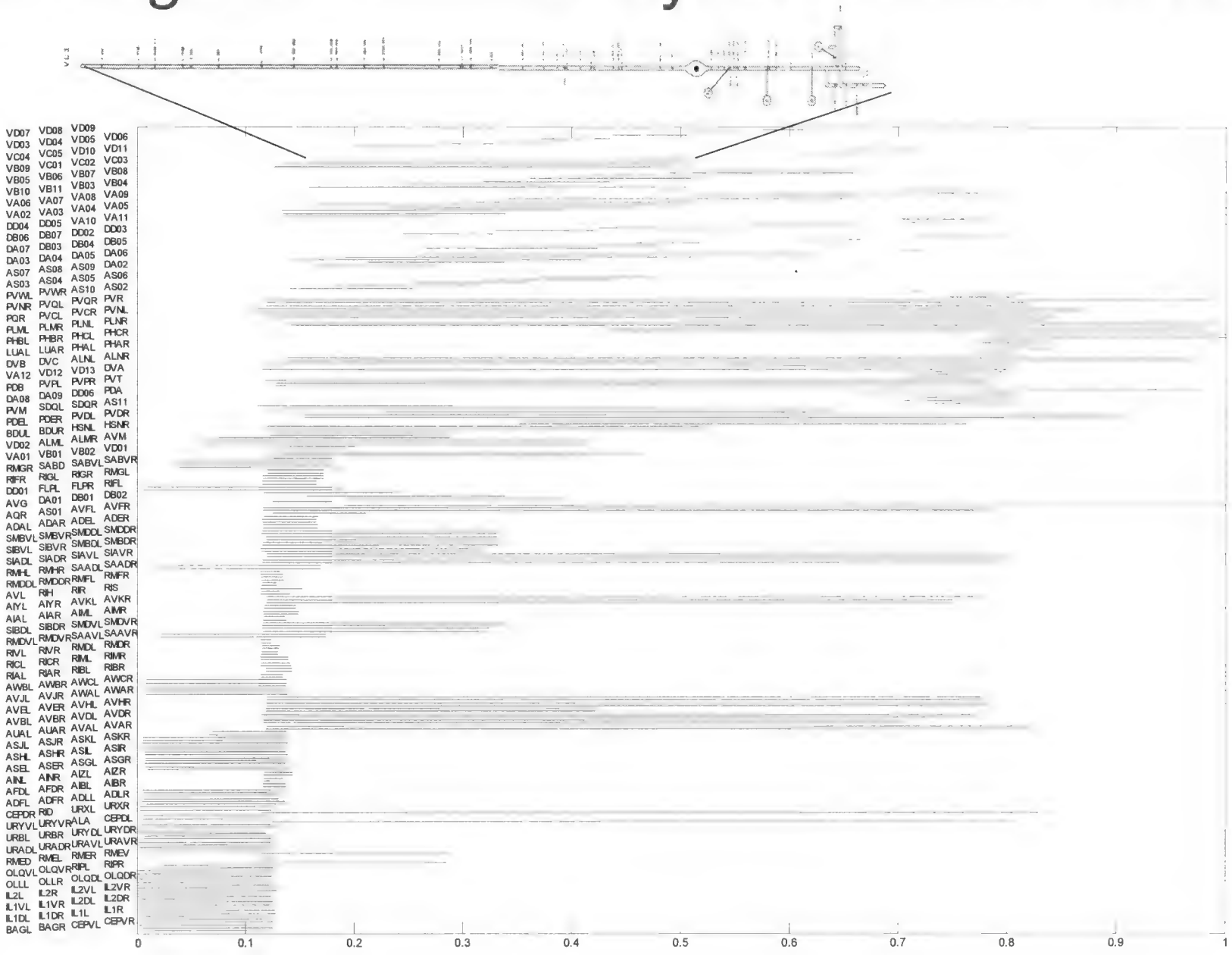
C. Elegans wiring diagram



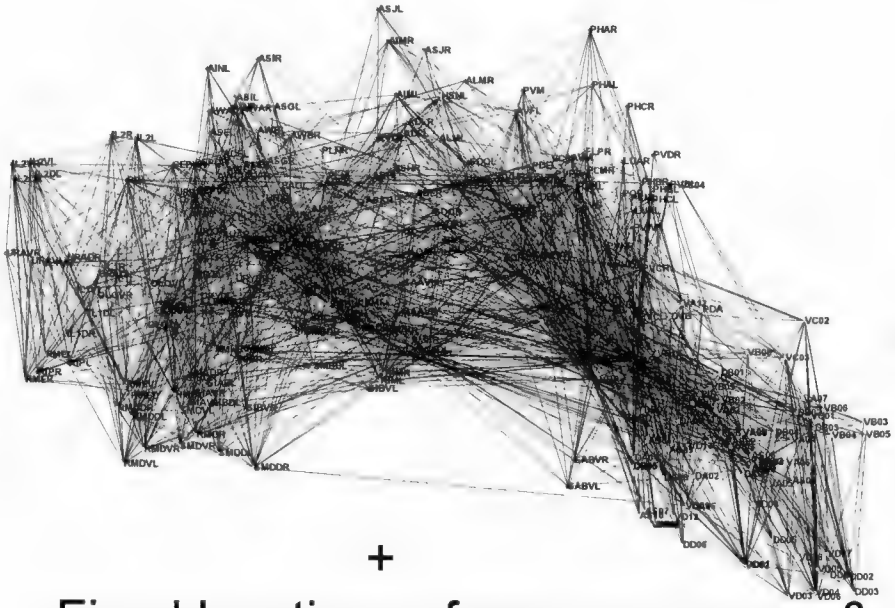
279 neurons, ~7000 synapses

Partially completed by J. White et al. (1986); Finalized by Chen, Hall & Chklovskii (2006)

C. elegans neuronal layout reduced to 1D



Can wiring diagram predict neuronal shape and layout?

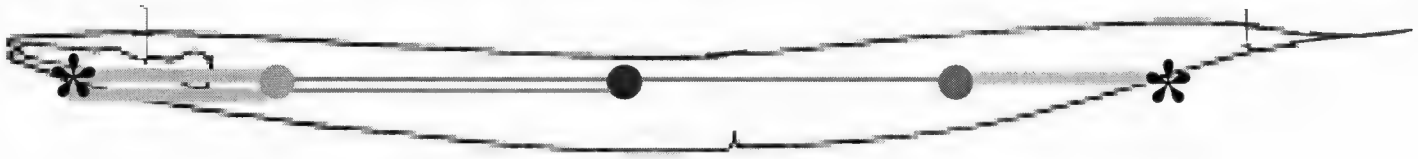


+

Fixed locations of sensory organs & muscles

?

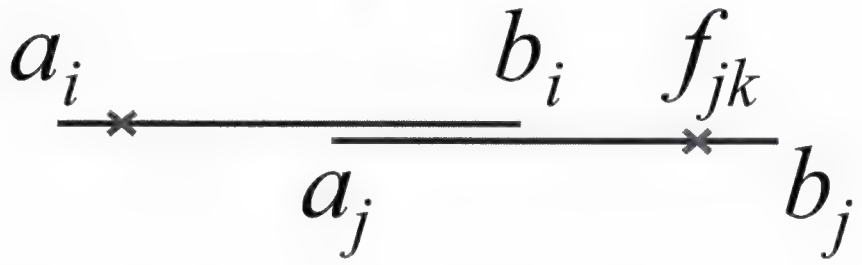
Method: Minimizing the cost of wiring



Optimization problem

- Minimize total wiring length:

$$L = \sum_i | b_i - a_i |$$



- Given connectivity constraints:

If neurons i and j are connected

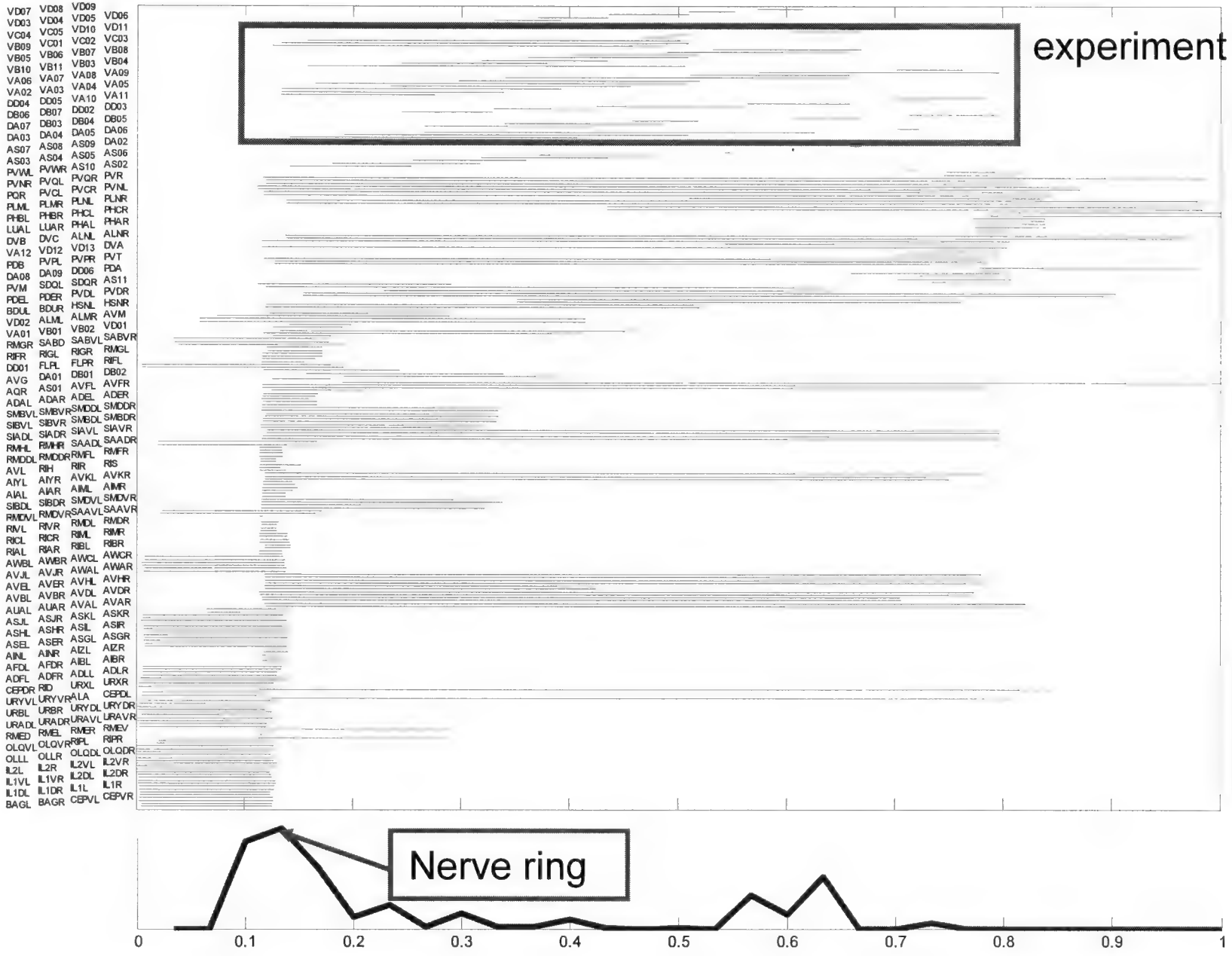
$$a_j \le b_i$$

$$a_i \le b_j$$

If neuron j has a sensory ending or a neuromuscular junction

$$a_j \le f_{jk} \le b_j$$

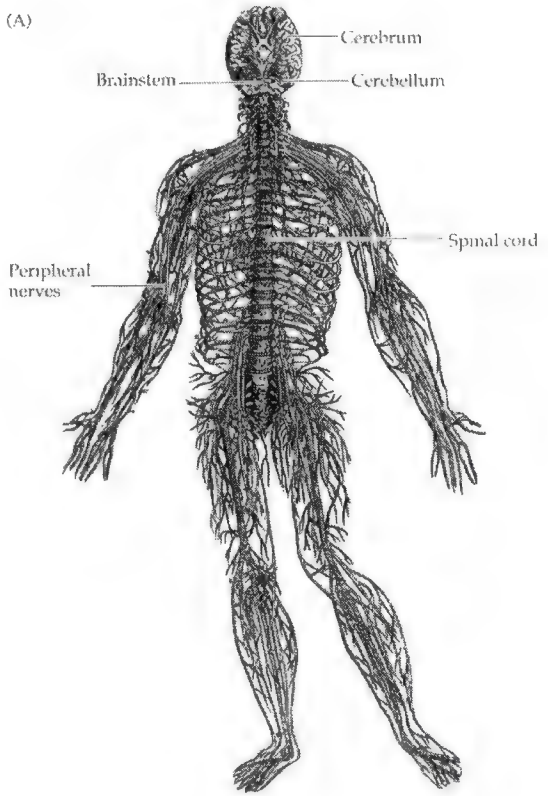
Optimized layout is close to actual



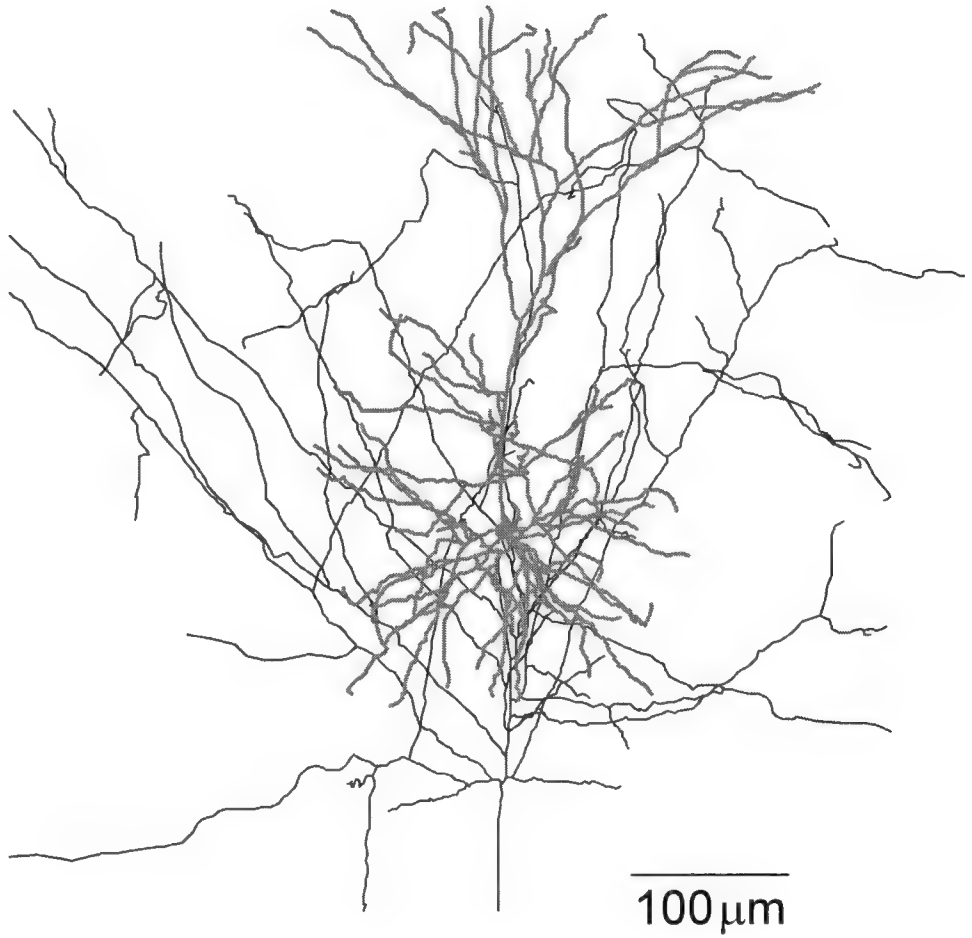
For given connectivity, wiring optimization largely predicts neuronal layout in invertebrates and long-range connections in vertebrates

Cherniak, Mitchison, Van Essen, Klyachko & Stevens, de Polavieja

10cm



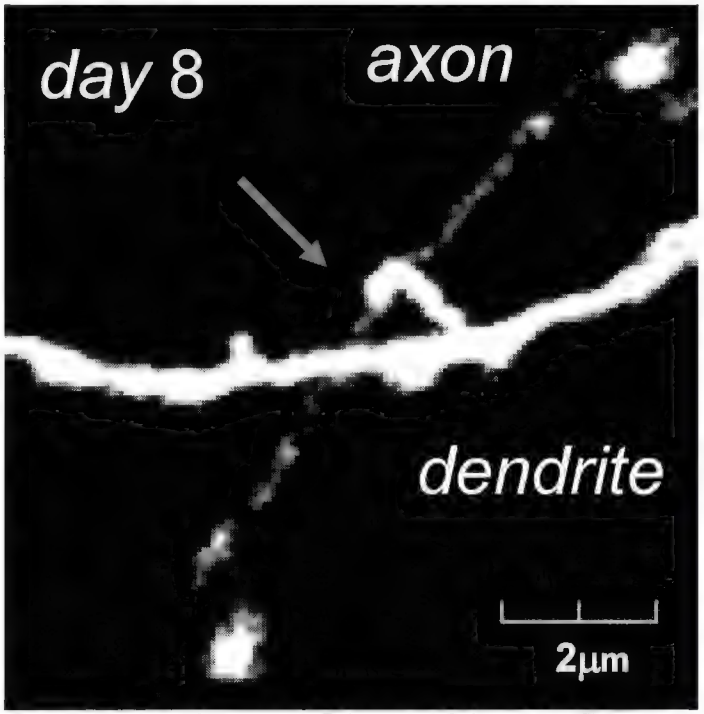
A typical neuron has millimeters of dendrites and centimeters of axons



Shepherd

Why do neurons have extensive axons and dendrites?

Synaptic connectivity in mammalian cortex can change with time



Trachtenberg, Chen, Knott, Feng, Sanes, Welker, Svoboda, 2002

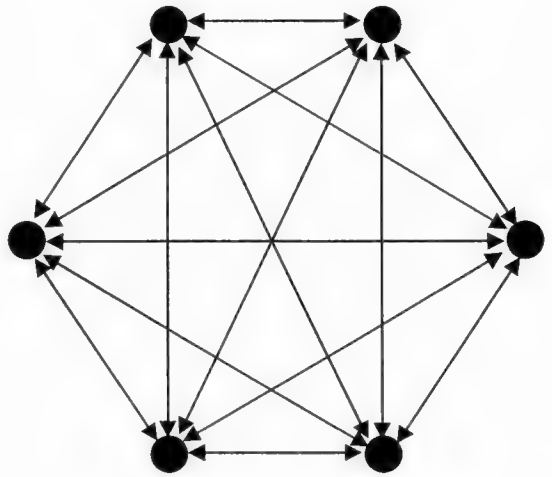
Potential synapse: proximity between a dendrite and an axon that can be bridged by a spine

Stepanyants, Hof & Chklovskii (2002)

Wiring problem

What is the smallest volume of the all-to-all connected network of N neurons with wires of diameter d ?

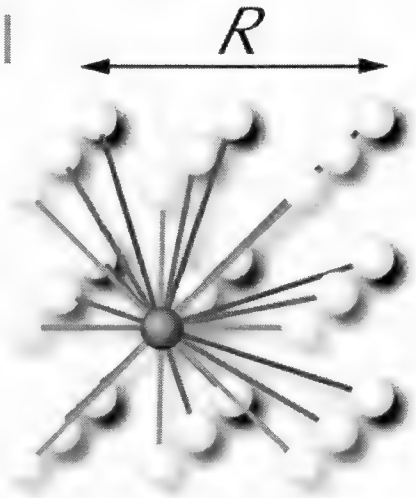
Example, $N = 6$:



Wiring designs for an all-to-all network

Point-to-point axons

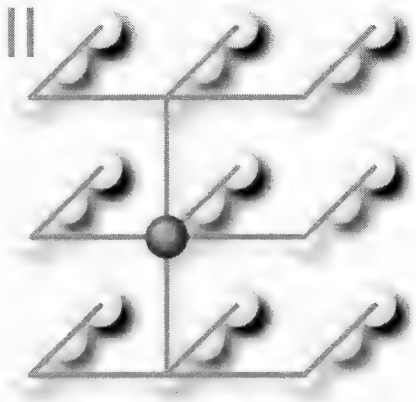
I



$R^3 \sim N^3 d^3$

Multi-terminal axons

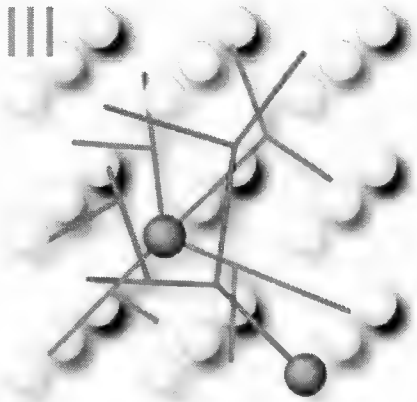
II



$R^3 \sim N^{5/2} d^3$

Multi-terminal axons & dendrites

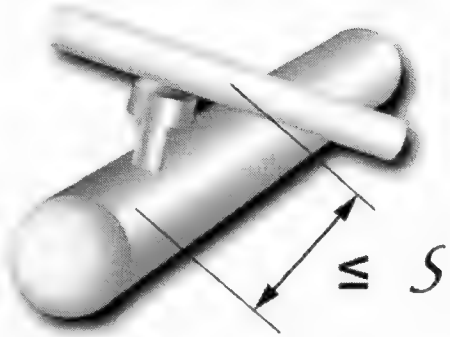
III



$R^3 \sim N^2 d^3$

Multi-termnal axons & spiny dendrites

IV

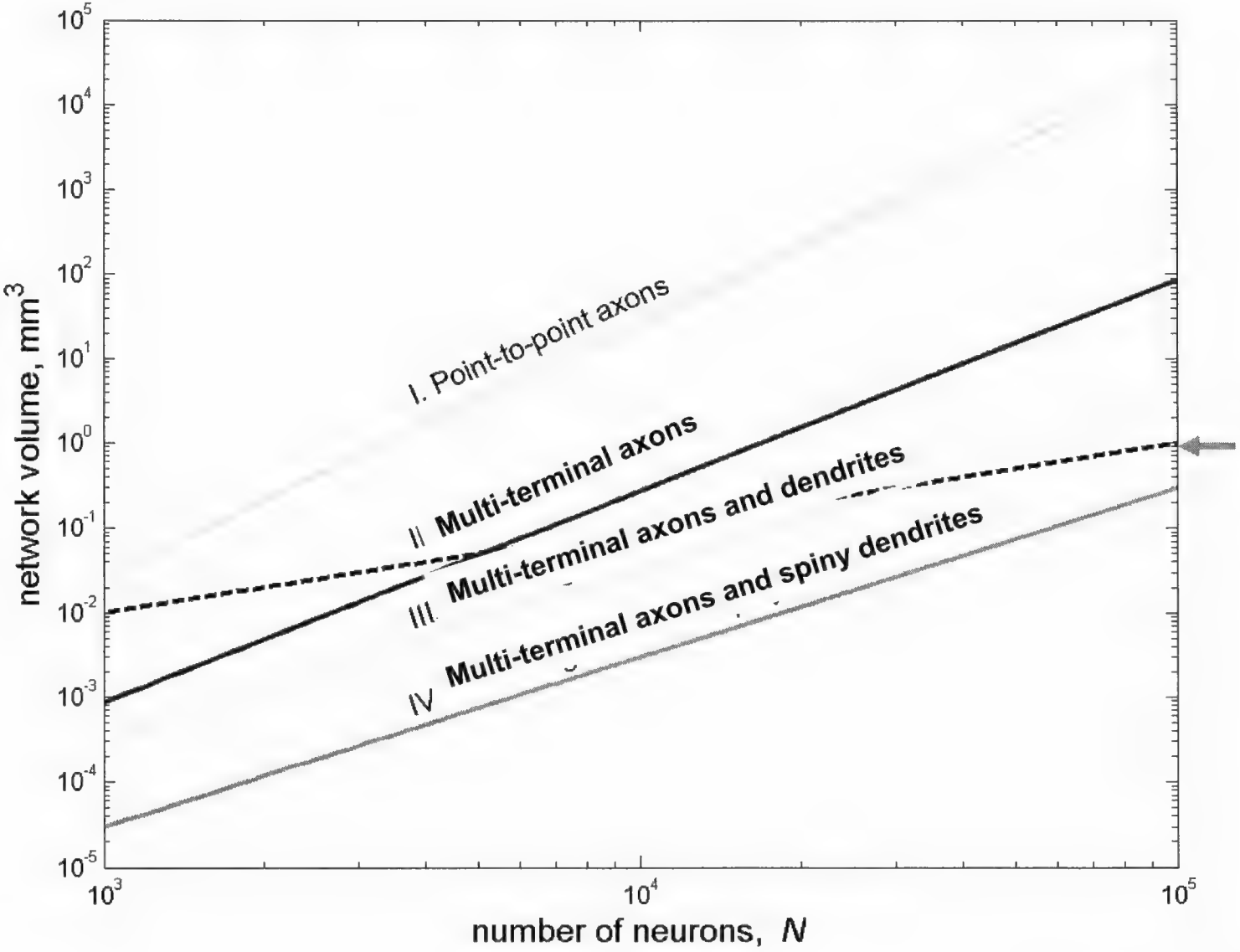


$R^3 \sim N^2 d^4/s$

Cortical column: $N=10^5$, $d=0.2\mu m$; $0.5\mu m$, $s=1.5\mu m$, $R=1mm$

$8{,}000 mm^3$ $30 mm^3$ $2 mm^3$ $0.6 mm^3$

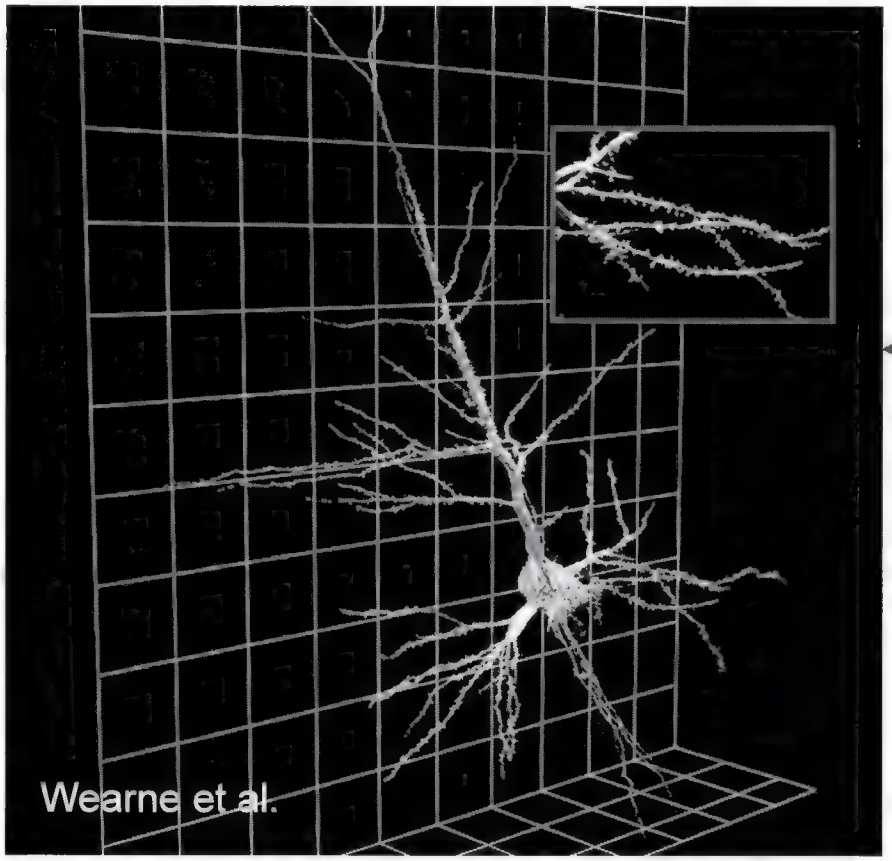
Network volume for different designs

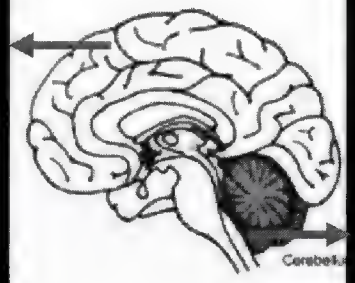


Chklovskii, 2004

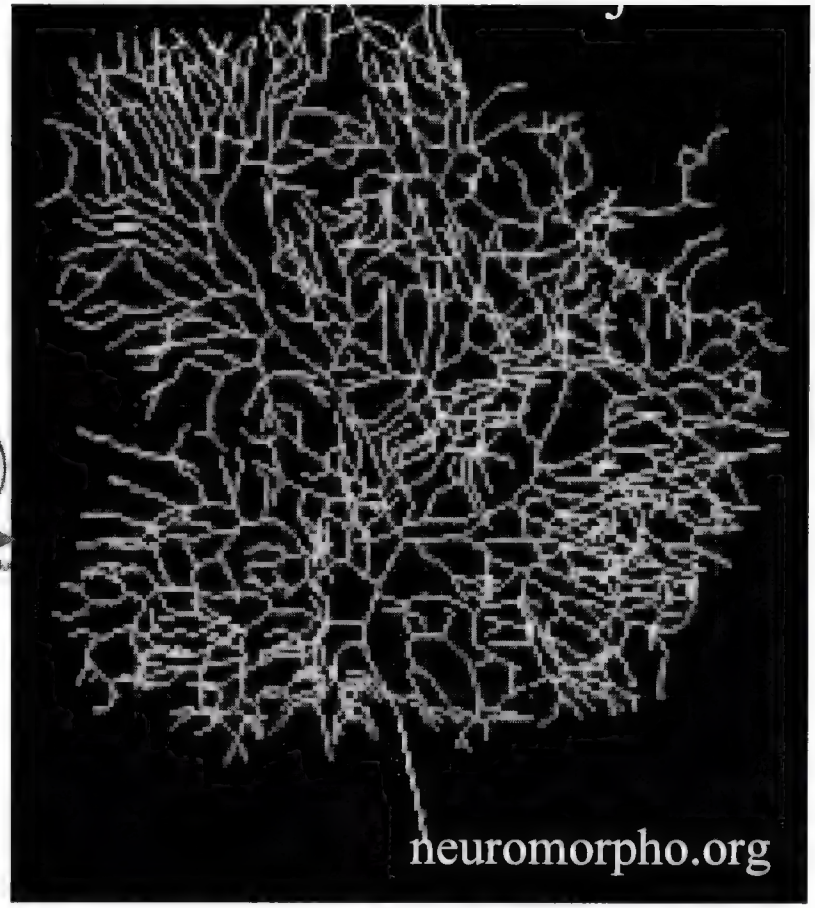
What determines the shape of dendritic arbors?

Pyramidal cell



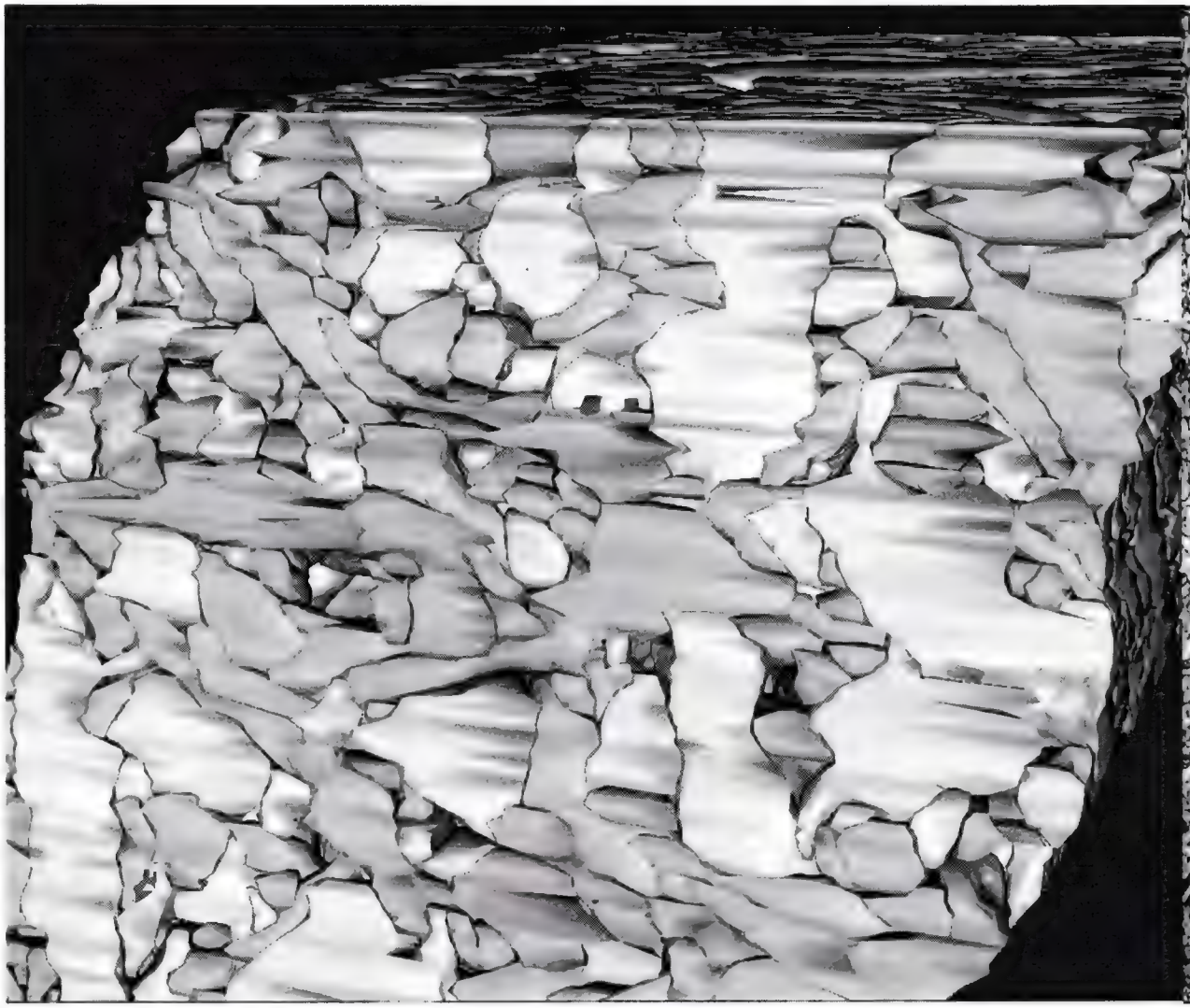


Purkinje cell



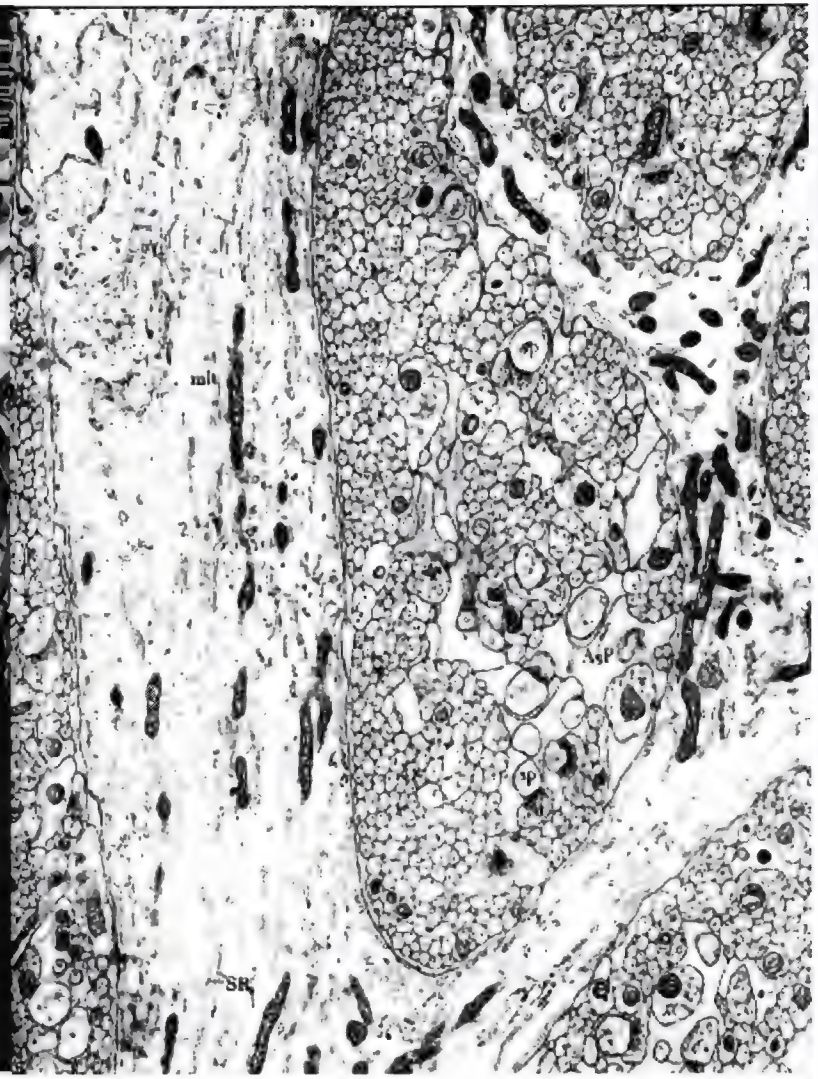
Arbors maximize the connectivity repertoire while minimizing wiring cost

Packing of axons and dendrites is space-filling



Hippocampus: axons in all directions

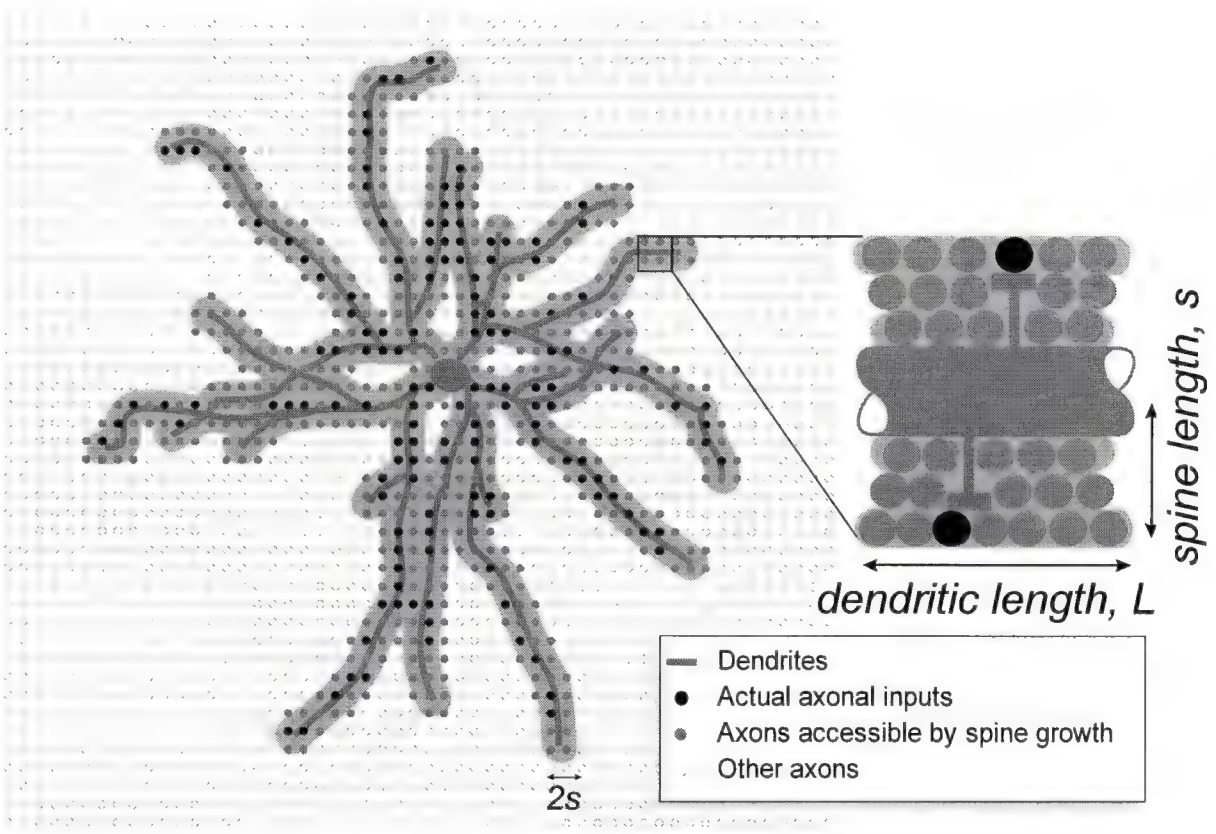
Mishchenko, Spacek, Mendenhall, Harris, Chklovskii



Cerebellum: axons orthogonal to dendrites

Sanford Palay

What determines the total length of dendrites? or The airport terminal theory



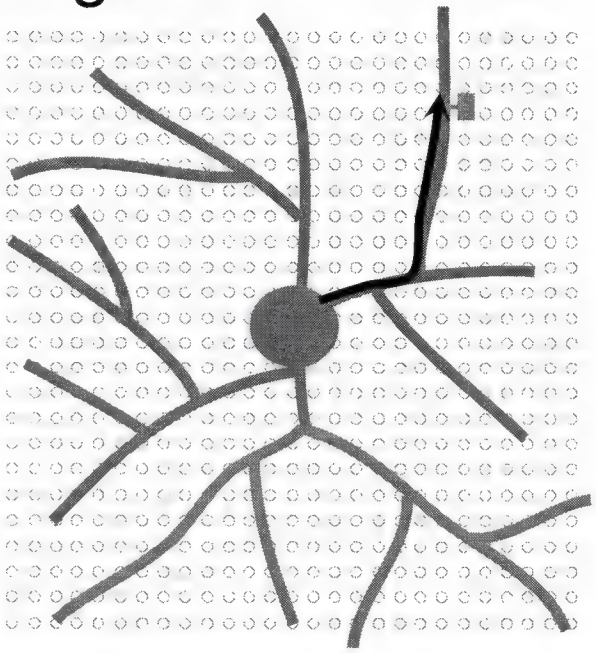
New York JFK airport

Total dendritic length, *L*, determines the number of potential synapses, *N*: $L \sim Nd^2/s$

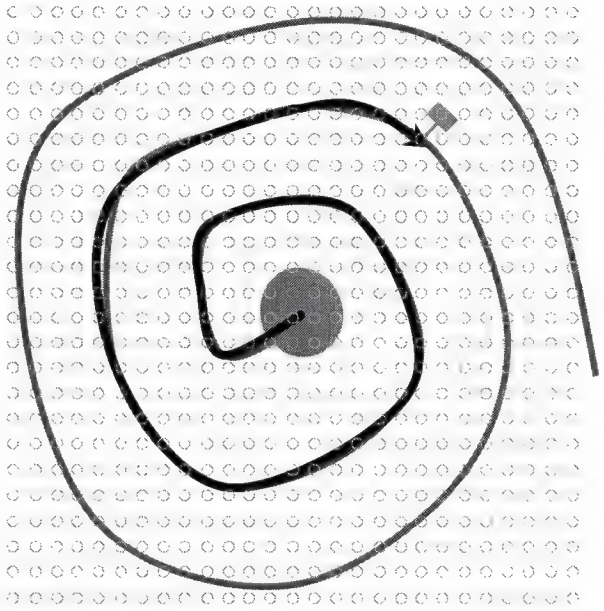
Pyramidal cells: $N = 10^5$ $d = 0.2\mu m$ $s = 1.5\mu m$ $\Longrightarrow$ $L = 2.6mm$

Why do dendrites branch?

Two dendrite designs with the same total length and number of potential synapses



Branching arbor



Non-branching arbor

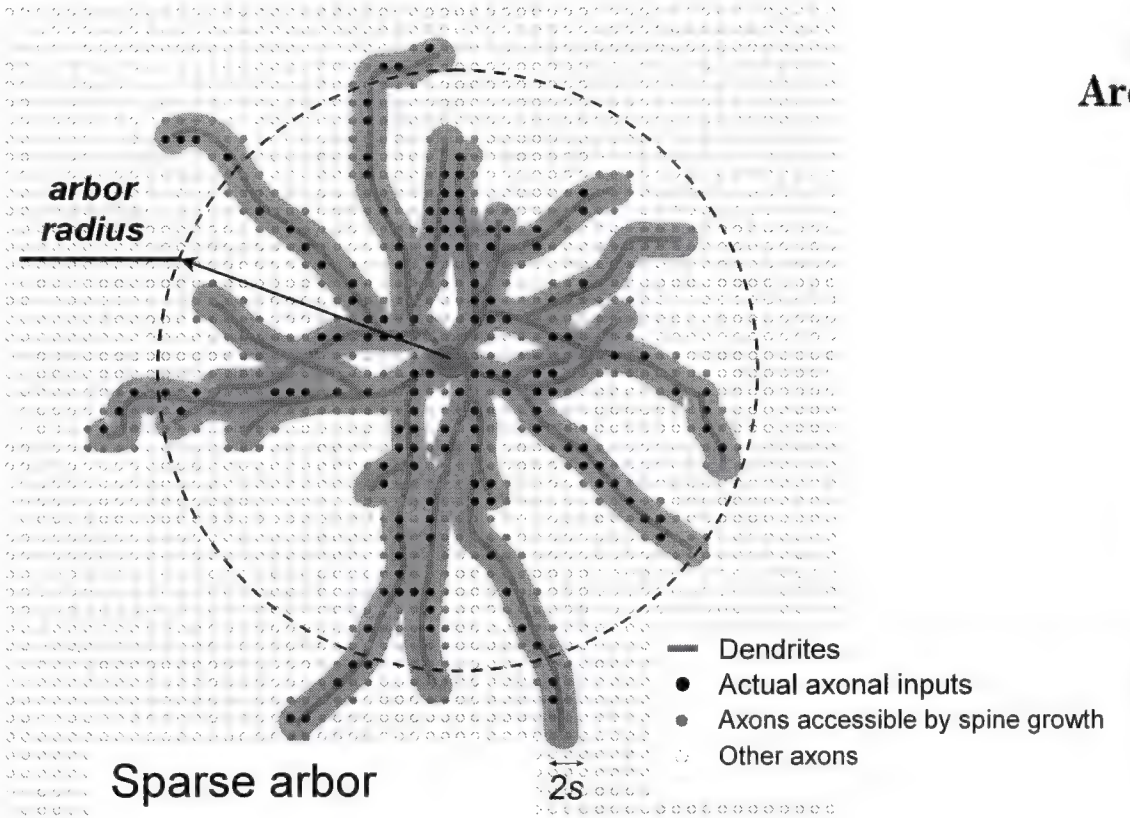
Chicago O'Hare airport

Branching dendrites have shorter path from synapses to cell body

Cuntz, Borst & Segev (2007), Wen & Chklovskii (2008)

Compact arbor minimizes path length for a given total dendritic length

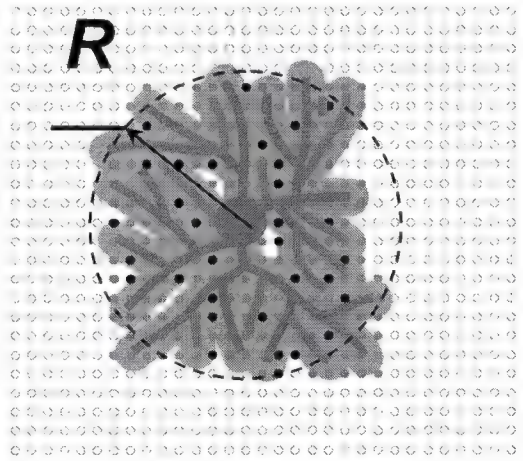
THE JOURNAL OF COMPARATIVE NEUROLOGY 361:479–490 (1995)



Sparse arbor

Retinal Neurons and Vessels Are Not Fractal But Space-Filling

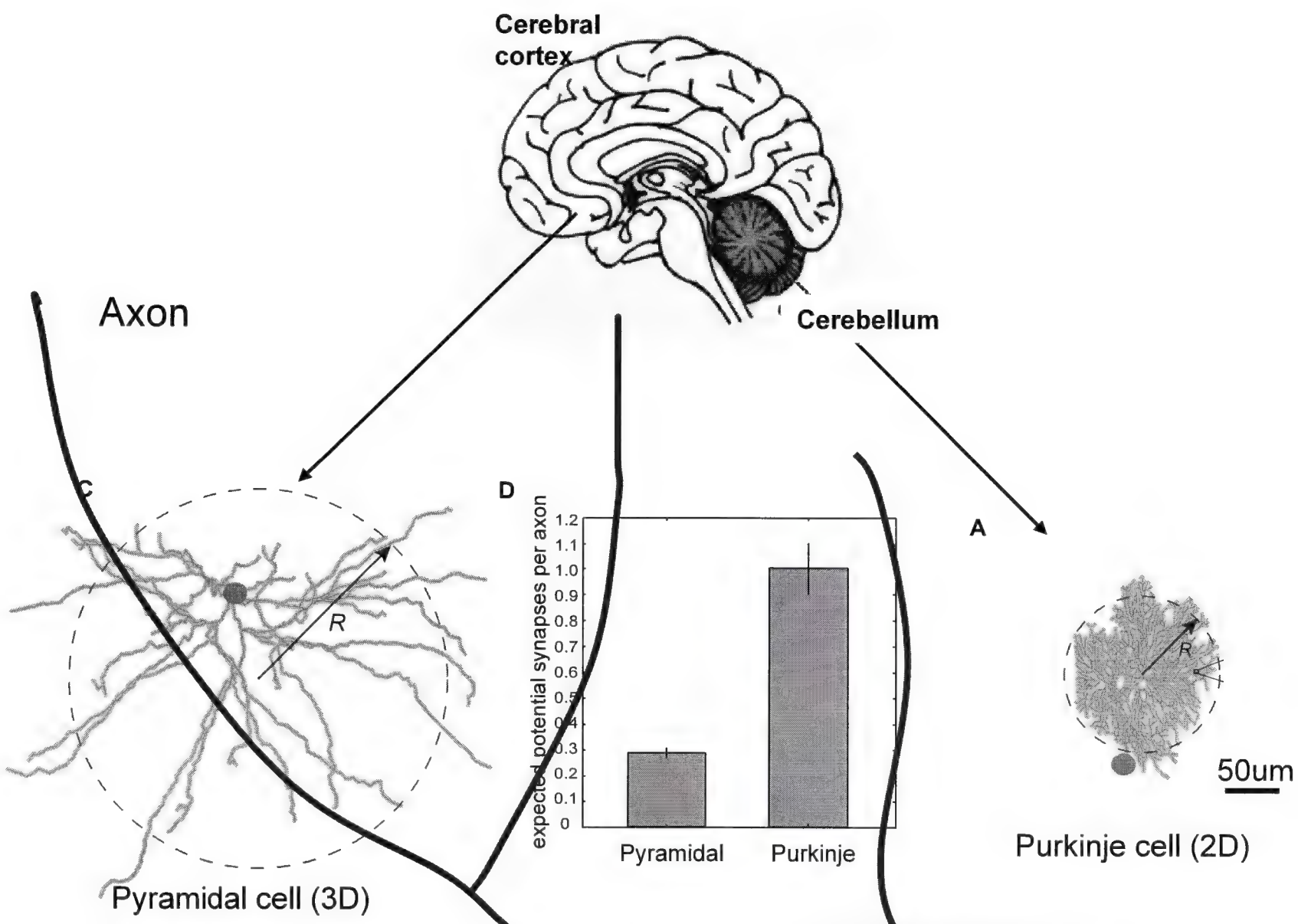
JOSEPH PANICO AND PETER STERLING
Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6058



$R \sim (Ls)^{1/2}$

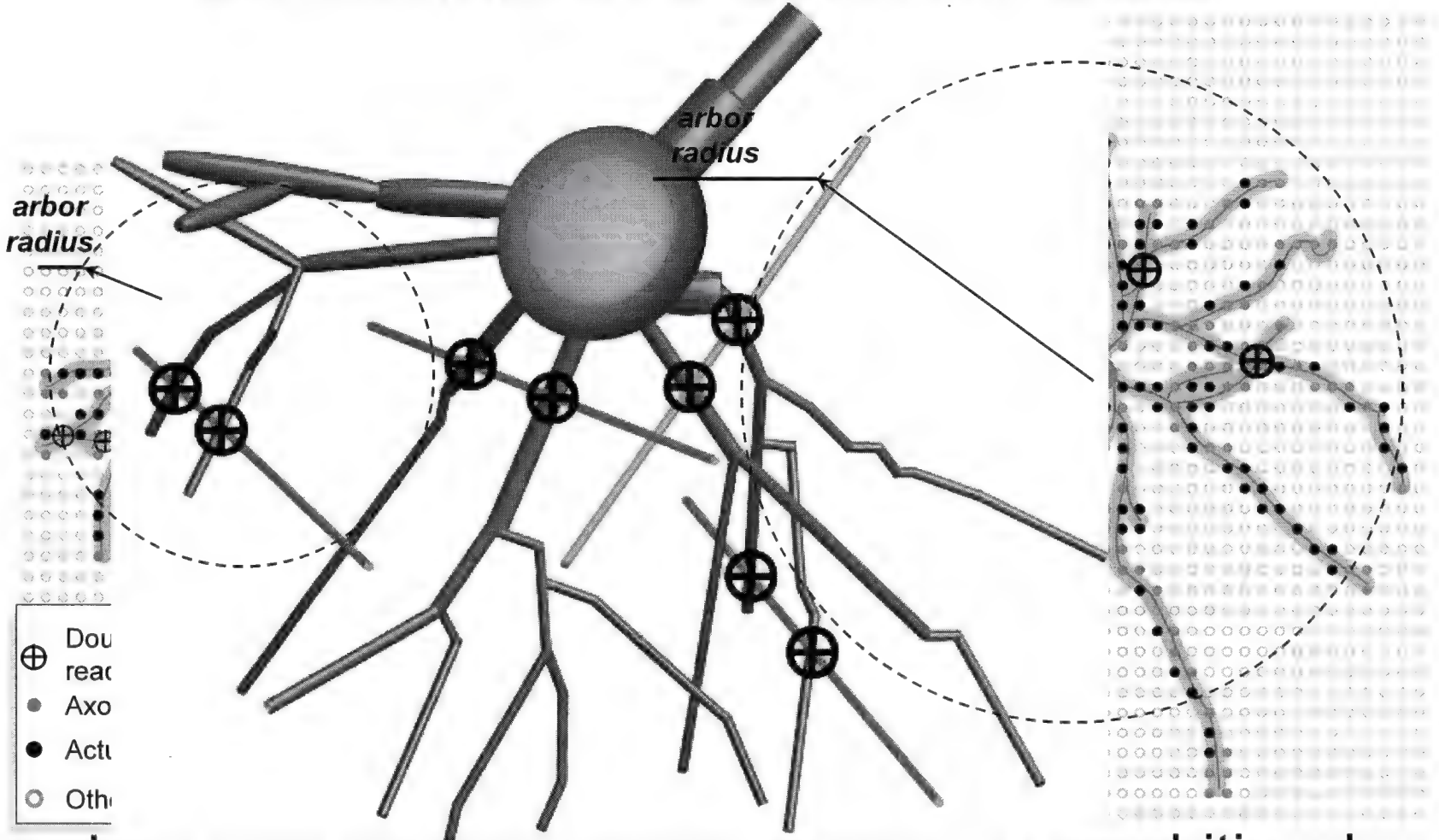
Compact arbor

For a given number of potential synapses, flat compact arbor minimizes wiring cost



Why are pyramidal cells sparse?

Double-hits reduce the number of axons accessible to a dendritic arbor



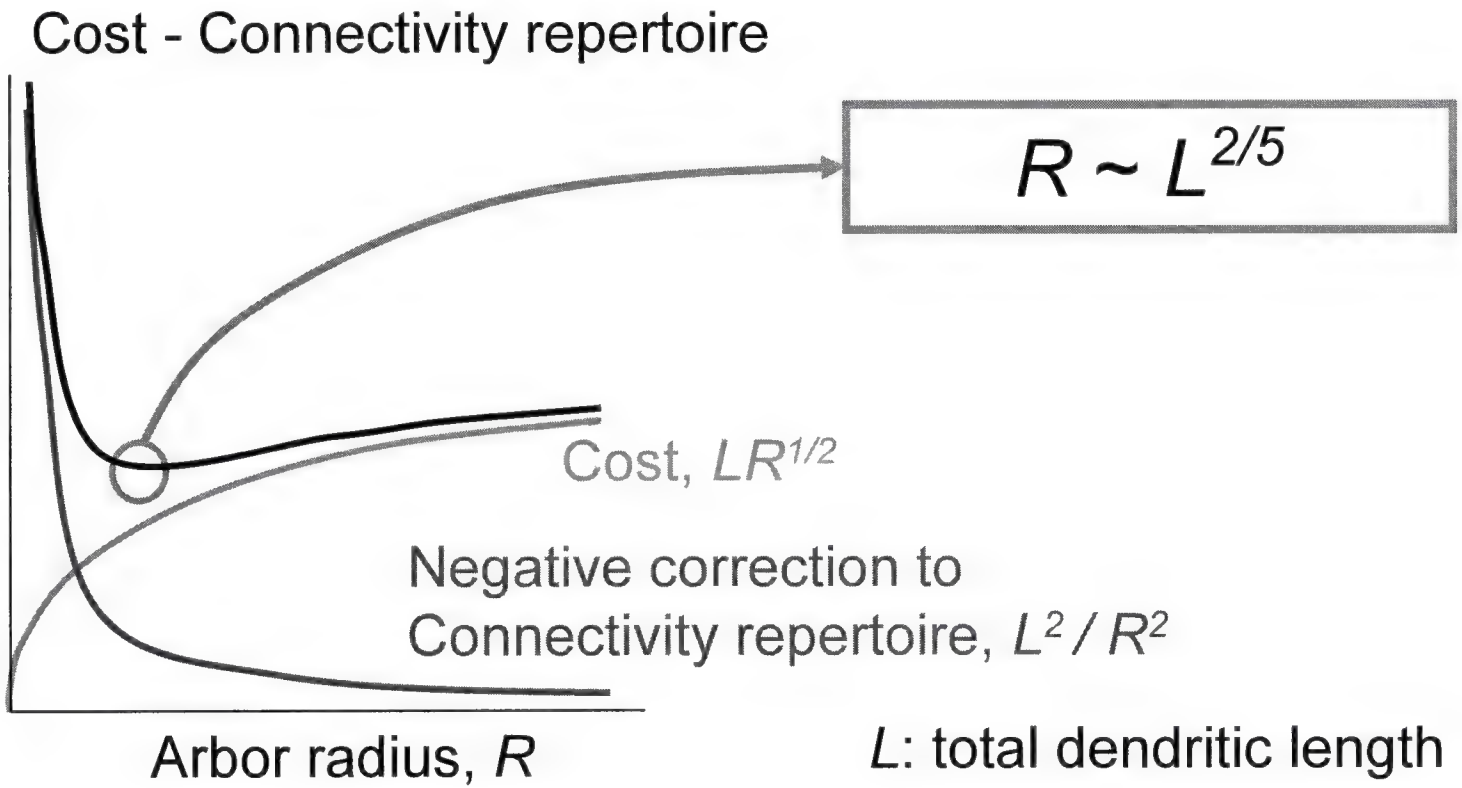
For a given total dendritic length, sparse 3D dendritic arbor has fewer double-hits, and, therefore, can access more axons

Statistical preference for sparse arbors, or self-repulsion

Samsonovich & Ascoli (2003)

Airport terminal with double-hits

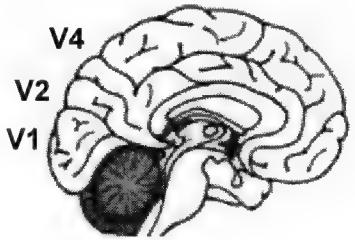
Arbor radius maximizes connectivity repertoire while minimizing cost = min{Cost - Connectivity repertoire}

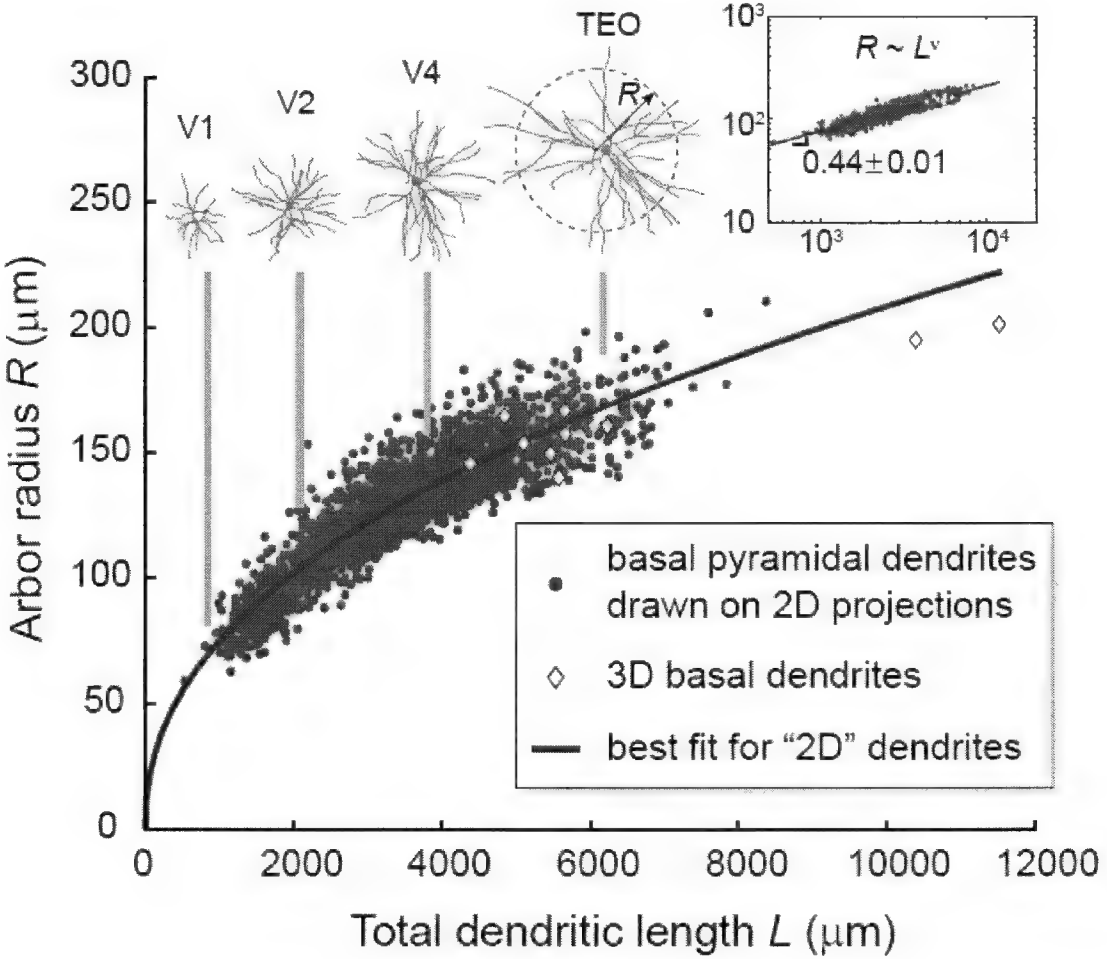


Free energy = Energy - Entropy

Scaling of basal dendrites of pyramidal cells

Cortical areas

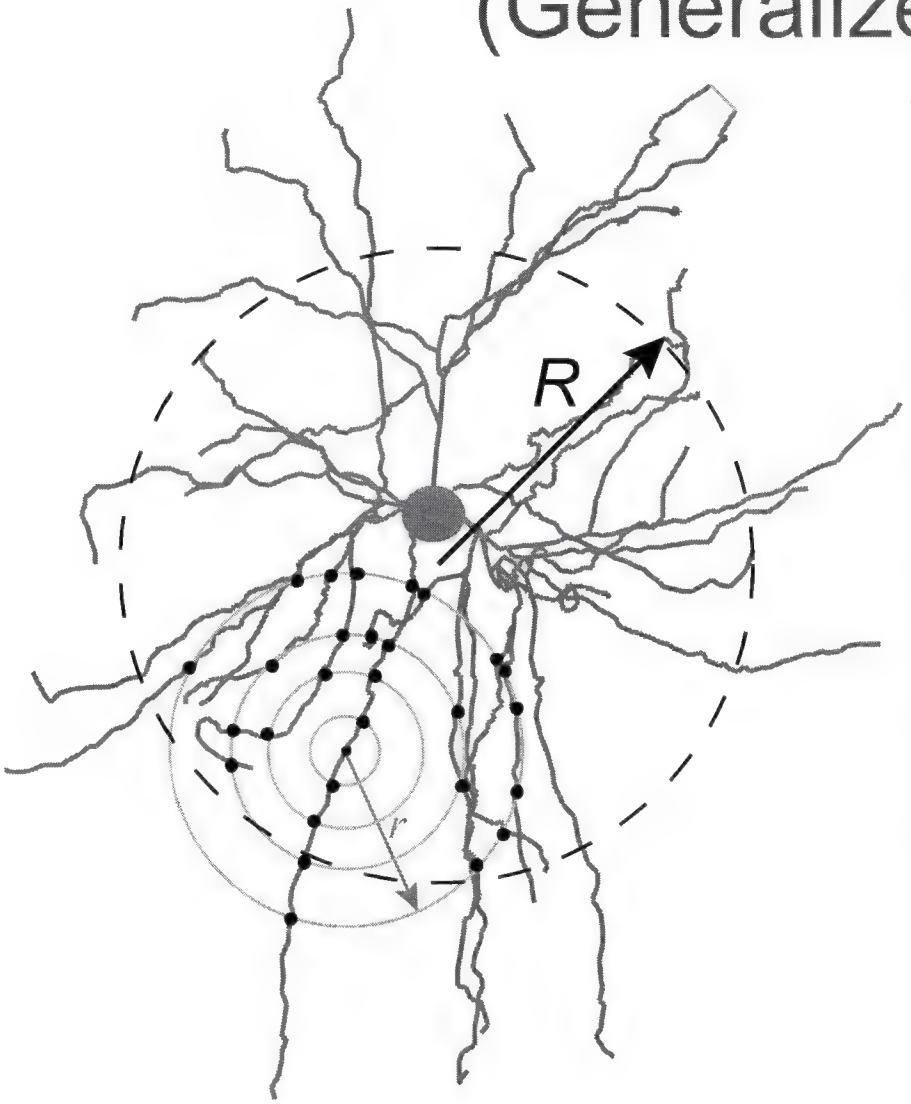


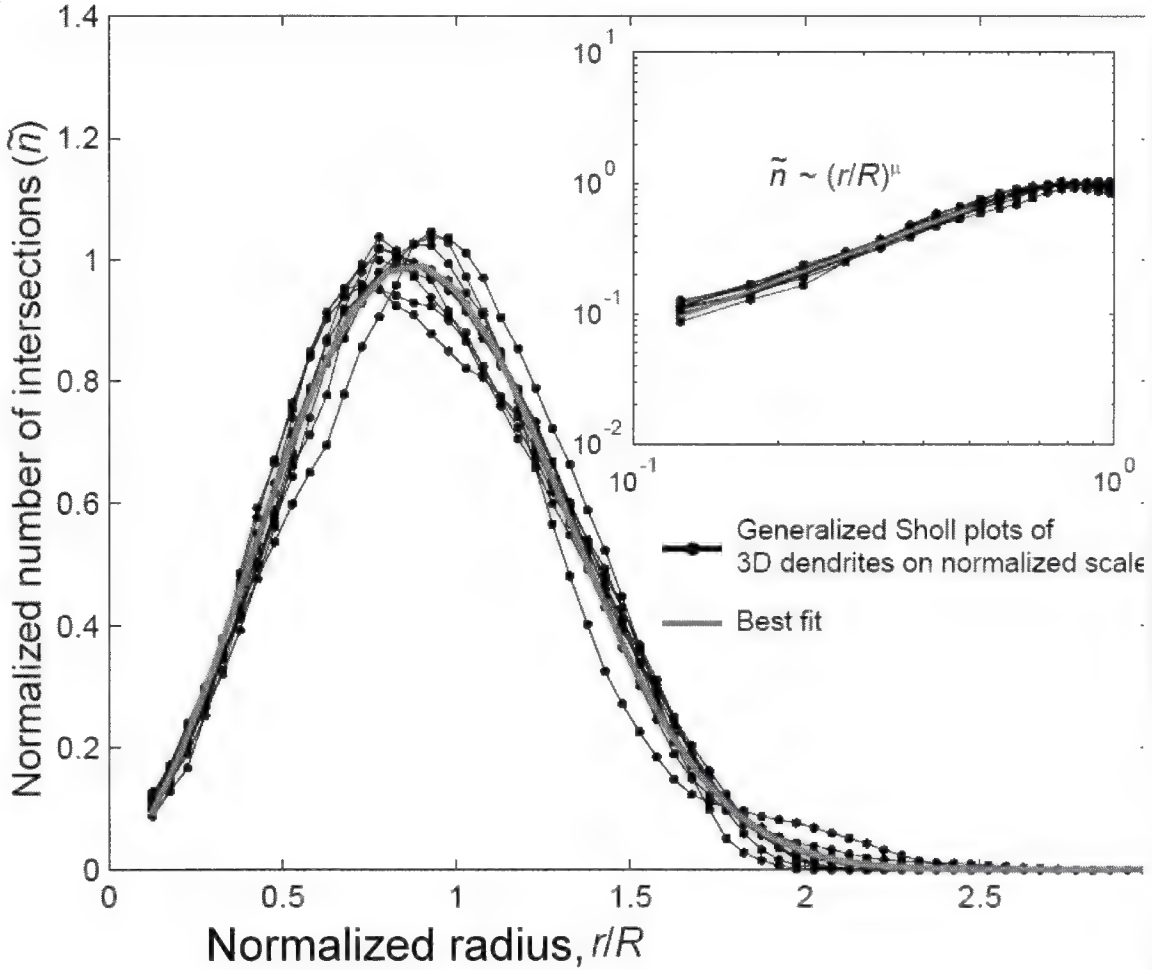


$R \sim L^{\nu}$ $\quad\quad$ $\nu = 0.44 \pm 0.01$

$\nu_{theory} = 0.4$

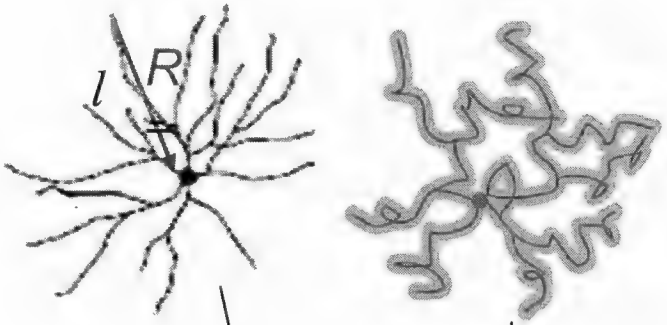
Dendritic correlation function (Generalized Sholl analysis)




Normalized number of intersections ($\tilde{n}$)
Normalized radius, r/R
$\tilde{n} \sim (r/R)^{\mu}$
Generalized Sholl plots of 3D dendrites on normalized scale
Best fit

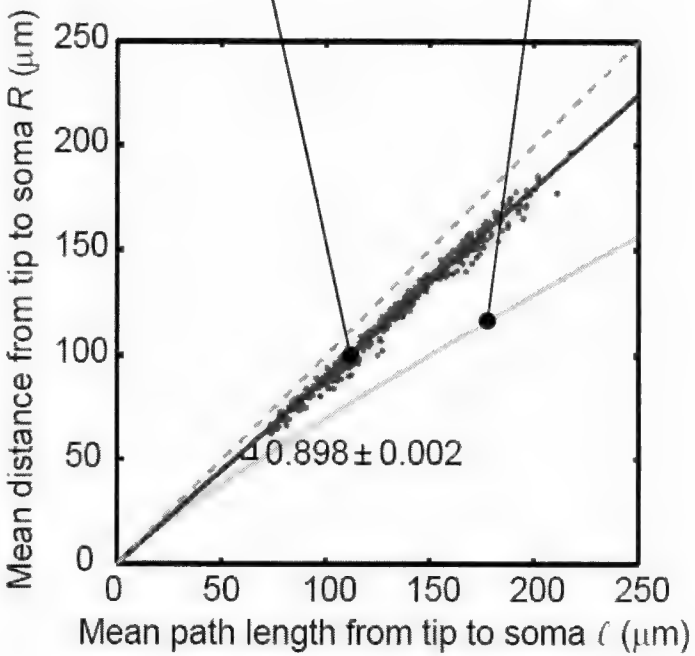
$$\tilde{n} = g(r / R)^{1/\nu - 1} \exp[-h(r / R)^{\gamma}]$$

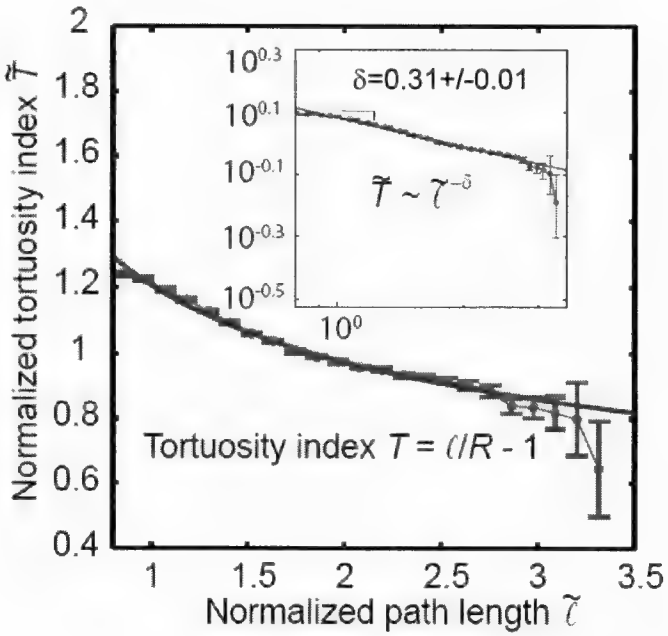
Dendrites are “stretched” by the path length cost



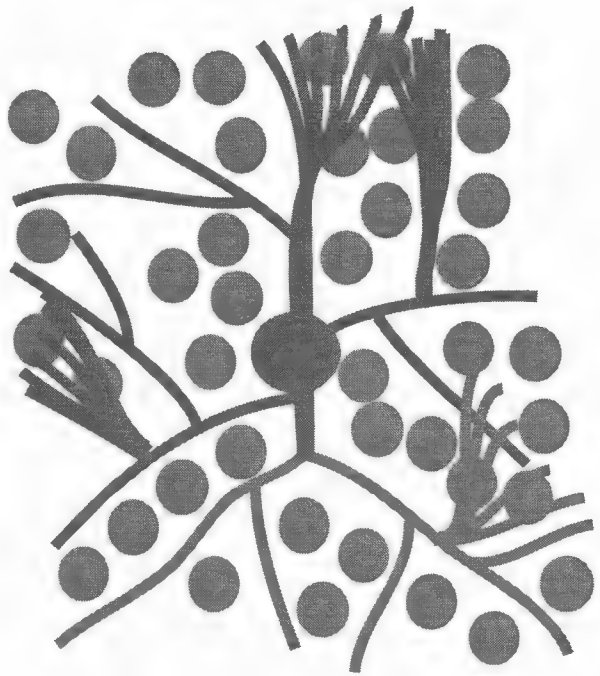
$$F \sim \frac{L^2}{R^2} - \frac{L}{a}\log(1 - R/\ell) + \beta L \ell^{\delta}$$

$$\frac{\ell}{R} - 1 = \frac{1}{\delta\beta\ell^{\delta} a}$$





Why aren't Purkinje dendrites sparse?



Branches of a flat (2D) arbor can be arranged to avoid double hits

Summary

- Minimizing wiring cost for given connectivity explains neuronal layout in inverterbrates and long-range connections in vertebrates

- Minimizing wiring cost while maximizing connectivity repertoire can account for the shape of mammalian neurons

Acknowledgements

Beth Chen (Cold Spring Harbor Laboratory)
Quan Wen (CSHL, Janelia, Harvard)
Armen Stepanyants (CSHL, Northeastern)
Alexander Grosberg (NYU)
Guy Elson (Queensland)

Rava Azeredo da Silveira, Woods Hole, August 2009

PROBLEM SET

Simple Stuff in Information Theory

PROBLEM 1: Entropy of a Spike Train vs. Entropy of a Spike Count

Consider a spike train of duration T, in which the neuron fires a spike with probability r per unit time. Suppose that the readout mechanism acts over a time scale τ, so that τ sets the resolution of the system. So imagine splitting the spike train duration, T, in a succession of short bins of duration τ each. In each bin, a spike is fired with probability $r\tau$ and no spikes are fired with probability $1 - r\tau$. (We assume that τ is small enough that the probability of two or more spikes per bin is negligible.)

(a) How many spikes are fired, on average, in the whole spike train? If T is sufficiently large, instantiations with more or less than the average number of spikes are very rare, so we can ignore them in calculating the entropy. What is, then, the total number of possible, distinct spike trains? Are these equiprobable?

(b) Combining the these considerations, write down an expression for the entropy of the spike train, H_{train}, in terms of r, τ, and T. Using Stirling's approximation of the factorial for large argument,

$$n! \approx \left(\frac{n}{e}\right)^n , \tag{1}$$

show that the entropy takes the simple form

$$H_{\text{train}} \approx -\frac{1}{\ln(2)} \frac{T}{\tau} \left[r\tau \ln(r\tau) + (1 - r\tau) \ln(1 - r\tau) \right] . \tag{2}$$

Give an intuitive explanation of this result. (*Hint*: Base your arguments on the fact that bins are independent from each other.)

(c) In the limit of low firing rate (which we have been assuming all along), $r\tau \ll 1$, show that this expression reduces to

$$H_{\text{train}} \approx -\frac{1}{\ln(2)} rT \left[\ln(r\tau) - 1 \right] . \tag{3}$$

What is the entropy per spike? Can it be larger than one? Explain why. Can you give an intuitive argument according to which we can expect $H_{\text{train}} \sim \ln(r\tau)$? (*Hint*: If the firing rate is r, a given spike is bound to be emitted in a window of time of order $1/r$.)

(d) It is instructive to compare the entropy of a spike train to the entropy of a spike count. For the comparison to make sense, we have to assume that

the mean spike count is identical to the mean number of spikes in the train, rT. What is the spike count probability, with mean rT, which maximizes the entropy? Show that the entropy of this distribution can be calculated as

$$H_{\text{count}} = \frac{rT}{\ln(2)} \ln\left(1 + \frac{1}{rT}\right) + \frac{1}{\ln(2)} \ln(1 + rT). \tag{4}$$

(e) Typically, $H_{\text{train}} > H_{\text{count}}$. Why? Examine now the (unusual) limit in which there are very few spikes in the entire train, $rT \ll 1$. Show that, in the limit, we have

$$H_{\text{count}} \approx -\frac{rT}{\ln(2)} [\ln(rT) - 1]. \tag{5}$$

Thus, in this limit, the spike count entropy starts looking like the spike train entropy. How do you make sense of this?

PROBLEM 2: Equivalent Expressions of the Mutual Information

Show that the mutual entropy can be written in several equivalent ways:

$$\begin{aligned} I &= H(s) + H(r) - H(s,r) && (6)\\ &= H(s) - H(s \mid r) && (7)\\ &= H(r) - H(r \mid s). && (8) \end{aligned}$$

Show that it can also be written as

$$I = H(s) - \left[\langle -\ln(P(r \mid s) P(s))\rangle - \left\langle -\ln\left(\sum_s P(r \mid s) P(s)\right)\right\rangle\right]. \tag{9}$$

Can you give an intuitive understanding of this form?

PROBLEM 3: "Histogram Equalization"

Consider a model of neural response that assumes additive variability, as in

$$r = m(s) + \delta, \tag{10}$$

where the neural response, r, is obtained as the sum of a deterministic function of the stimulus, $m(s)$ (often referred to as the 'tuning curve'), plus a noise, δ. We assume that the stimulus varies between $s_{\min}$ and $s_{\max}$, and that the response varies between 0 and $r_{\max}$.

In the case of additive noise, the mutual information is simply the difference between the entropy of the response and that of the noise,

$$I = H(r) - H(\delta). \tag{11}$$

Here, we look for the form of the processing, *i. e.*, the form of $m(s)$, that maximizes I.

If we assume that the noise is small with respect to the range of the response, we can assume that the entropy of the response is dominated by the entropy of the stimulus. Thus, maximizing I amounts to maximizing $H(m)$, where m is a random variable distributed like $m(s)$ is.

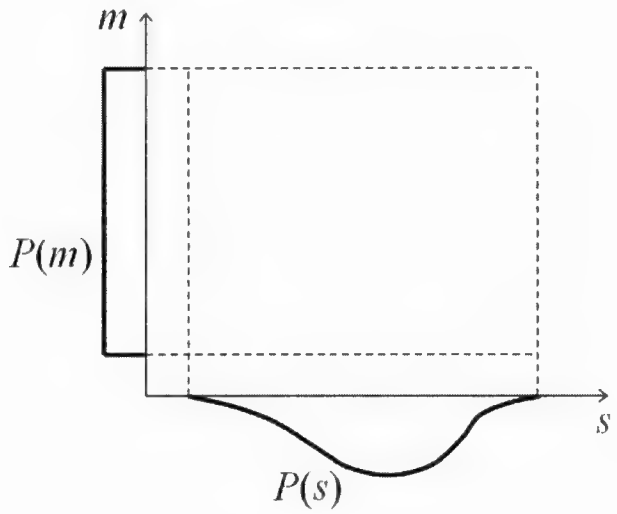
(a) Show that the probability distribution that maximizes $H(m)$ is a uniform distribution, $P(m) = \text{constant}$.

(b) Show that this implies that the mutual information is maximized by the tuning curve

$$m(s) = r_{\max} \int_{s_{\min}}^{s} d\tilde{s} dP(\tilde{s}). \tag{12}$$

(*Hint*: From the definition of M, we have $P(m)\, dm = P(s)\, ds$.)

(b) Give a graphical interpretation of this result. If you like, you can use and complete the plot below.



PROBLEM 4: Maximizing the Mutual Information in a Linear Neuron

We consider a noisy neuron which responds linearly to a set of N stimuli, as

$$r = \sum_{i=1}^{N} w_i s_i + \delta, \tag{13}$$

or, in vector notation

$$r = w^T s + \delta, \tag{14}$$

where w is the vector of linear weights. It cannot be arbitrarily large, and we assume that its norm is bounded by some number (say, 1):

$$|w|^2 = \sum_{i=1}^{N} w_i^2 \leq 1. \tag{15}$$

(We could have picked more or less any form of the norm, but we chose the Euclidean norm to fix the ideas.) We assume that both the stimuli and the noise are distributed according to zero-mean Gaussians, so that their probabilities read

$$P(s) = \frac{1}{\sqrt{(2\pi)^N \, DetC}} \exp\left(-\frac{1}{2} s^T C^{-1} s\right) \tag{16}$$

and

$$P(\delta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{\delta^2}{\sigma^2}\right). \tag{17}$$

Thus, the covariances of the stimuli are obtained as

$$\langle s_i s_j \rangle \equiv C_{ij}. \tag{18}$$

(a) Calculate the probability distribution of the response, $P(r)$.
(b) Calculate the mutual information, I.
(c) For what choice of weight vector, w^*, is the mutual information maximized? (*Hint*: You may find it helpful to express w as a linear combination of orthonormal eingenvectors of the matrix C, $e^{(i)}$ with $i = 1, ..., N$, as

$$w = \sum_{i=1}^{N} a_i e^{(i)}, \tag{19}$$

where the vectors $e^{(i)}$ satisfy

$$e^{(i)T} e^{(j)} = \delta_{ij}.) \tag{20}$$

(d) Can you make sense of your result for w^*? Explain in a few sentences.

Rava Azeredo da Silveira, Woods Hole, August 2009

PROBLEM SET

Fisher Information of Correlated Neurons

The aim of this problem set is to become familiar with the Fisher Information approach to populations of correlated neurons.

We consider a population of N neurons, indexed by $i = 0, \ldots, N-1$, which respond to a continuous stimulus, s. For the sake of simplicity, we consider periodic boundary conditions. (Think, for example, of an angle stimulus and the corresponding orientation tuning often discussed in the literature.) Each neuron responds to a presentation of the stimulus by firing a number of spikes,

$$r_i(s) = m_i(s) + \eta_i. \tag{1}$$

The tuning curve of neuron i, $m_i(s)$, is the mean number of spikes fired in response to the stimulus presentation. We do not specify the shape of the tuning curves, but think of them as symmetric bell-shaped curves; the peak of the curve $m_i(s)$ defines the stimulus preference of neuron i. Furthermore, we assume that all the tuning curves have the same shape (they differ only by the location of their peak in stimulus space), with width W, and are uniformly distributed along stimulus space. Thus, if we call Δ the 'distance' between adjacent tuning curves in stimulus space, the stimulus space spans a total 'length' $N\Delta$. Below, we shall assume that tuning is 'broad', *i. e.*, for any given value of s a significant fraction of the population responds. In other words, $N\Delta/W \approx \mathcal{O}(1)$. (Note that the case in which each tuning curve 'covers' only a small number of neurons, $W \approx \mathcal{O}(\Delta)$, is less interesting as it is akin to single-cell coding.)

The second term on the right-hand-side, η_i, is a random variable with vanishing mean,

$$\langle \eta_i \rangle = 0, \tag{2}$$

and covariance

$$\langle \eta_i \eta_j \rangle \equiv C_{ij} = C_{|i-j|}. \tag{3}$$

Throughout, we assume that η_i is independent of the stimulus. We shall consider three qualitatively different cases:

- *vanishing correlations (uncorrelated neurons)*:

$$C_{|i-j|} = A\delta_{ij}; \tag{4}$$

- *spatially decaying correlations*:

$$C_{|i-j|} = A\delta_{ij} + B\rho^{|i-j|}(1-\delta_{ij}) \qquad \text{with} \qquad \rho \equiv e^{-\Delta/L}; \tag{5}$$

- *uniform correlations*:

$$C_{|i-j|} = A\delta_{ij} + B\left(1 - \delta_{ij}\right). \tag{6}$$

In the intermediate case the correlation length, L, measures the extent of correlations in stimulus space. We assume that it 'covers' a significant fraction of the population, *i. e.*, $N\Delta/L \approx \mathcal{O}(1)$. (We make this choice because the case with short-range correlations, $L \approx \mathcal{O}(\Delta)$, is akin to the uncorrelated case.) Note that the intermediate case reduces to the uncorrelated case when $\rho \to 0$ and to the uniform case when $\rho \to 1$.

(1) General expression of the Fisher information.

Recall the definition of the Fisher information,

$$I_F = \left\langle -\frac{\partial^2 \ln P\left(r \mid s\right)}{\partial s^2} \right\rangle_{P(r|s)}, \tag{7}$$

and assume that the random part of the response is distributed as a Gaussian,

$$P\left(r \mid s\right) = \frac{1}{\sqrt{\left(2\pi\right)^N DetC}} \exp\left(-\frac{1}{2} \sum_{i,j=0}^{N-1} \left[r_i - m_i\left(s\right)\right] \left(C^{-1}\right)_{ij} \left[r_j - m_j\left(s\right)\right] \right). \tag{8}$$

Show that the Fisher information can be written as

$$I_F = \sum_{i,j=0}^{N-1} \frac{\partial m_i}{\partial s} \left(C^{-1}\right)_{ij} \frac{\partial m_j}{\partial s} \tag{9}$$

or, in matrix notation,

$$I_F = \frac{\partial m}{\partial s} C^{-1} \frac{\partial m}{\partial s}. \tag{10}$$

(2) Expression of the Fisher information in terms of correlation eigenmodes.

Here we diagonalize the covariance matrix in order to obtain a simpler expression of the Fisher information, from which we shall extract some conceptual conclusions. Let the vectors

$$v^{(n)} = \begin{pmatrix} v_0^{(n)} \\ \vdots \\ v_{N-1}^{(n)} \end{pmatrix}, \tag{11}$$

with $n = 1, ..., N$, be the N orthonormal eigenvectors of the covariance matrix, C. Recall that the matrix V, made up of the eigenvectors,

$$V = \begin{pmatrix} v^{(0)} & \cdots & v^{(N-1)} \end{pmatrix} = \begin{pmatrix} v_0^{(0)} & \cdots & v_{N-1}^{(0)} \\ \vdots & & \vdots \\ v_{N-1}^{(0)} & \cdots & v_{N-1}^{(N-1)} \end{pmatrix}, \tag{12}$$

diagonalizes the covariance matrix:

$$V^T C V = \Lambda \equiv \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{N-1} \end{pmatrix}. \tag{13}$$

Here, λ_n is the eigenvalue that corresponds to eigenvector $v^{(n)}$.

Show that the inverse covariance matrix can be obtained as

$$C^{-1} = V \Lambda^{-1} V^T. \tag{14}$$

Using this result, show that the Fisher information can be written as

$$I_F = \sum_{n=0}^{N-1} \frac{M_n^2}{\lambda_n}, \tag{15}$$

where

$$M_n \equiv v^{(n)T} \frac{\partial m}{\partial s} = \sum_{i=0}^{N-1} v_i^{(n)} \frac{\partial m_i}{\partial s}. \tag{16}$$

(3) Diagonalization of the covariance matrix.

The covariance matrix, C, is an example of a circulant matrix, in which any line is obtained by shifting the previous line by one element:

$$C = \begin{pmatrix} & & & & \vdots & & & & \\ C_{-\frac{N-1}{2}+1} & \cdots & C_{-1} & C_0 & C_1 & \cdots & C_{\frac{N-1}{2}-1} & C_{\frac{N-1}{2}} & C_{-\frac{N-1}{2}} \\ C_{-\frac{N-1}{2}} & C_{-\frac{N-1}{2}+1} & \cdots & C_{-1} & C_0 & C_1 & \cdots & C_{\frac{N-1}{2}-1} & C_{\frac{N-1}{2}} \\ C_{\frac{N-1}{2}} & C_{-\frac{N-1}{2}} & C_{-\frac{N-1}{2}+1} & \cdots & C_{-1} & C_0 & C_1 & \cdots & C_{\frac{N-1}{2}-1} \\ & & & & \vdots & & & & \end{pmatrix}. \tag{17}$$

(Recall: $C_{ij} = C_{|i-j|}$.) Here and in the remainder we assume that N is odd.

Show that the normalized eigenvectors of a circulant matrix have the form

$$v^{(n)} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ z_n \\ z_n^2 \\ \vdots \\ z_n^{N-2} \\ z_n^{N-1} \end{pmatrix}, \tag{18}$$

where z_n is the n^{th} root of unity:

$$z_n = e^{i2\pi \frac{n}{N}}. \tag{19}$$

Show that the corresponding eigenvalue, of the covariance matrix, is

$$\lambda_n = C_0 + \sum_{j=1}^{\frac{N-1}{2}} C_j \left(z_n^j + z_n^{-j} \right), \tag{20}$$

with $n = 0, ..., N-1$.

This expression can be evaluated for each of the three correlation cases. Show the following.

- In the *uncorrelated case*, all eigenvalues are identical and equal to A. (*Hint*: We know this to be true by definition, but we can check the formula by inserting the values of the correlation matrix elements in the above sum.)
- In the *spatially decaying case*, we obtain

$$\lambda_n = A + 2B\rho \frac{\cos(2\pi n/N) - \rho - (-1)^n \rho^{(N-1)/2} (1-\rho) \cos(\pi n/N)}{1 - 2\rho \cos(2\pi n/N) + \rho^2}. \tag{21}$$

  (*Hint*: Sum the geometric series in the above sum, and recall that the eigenvalues must be real.)

- In the *uniform case*, we obtain

$$\lambda_n = \begin{cases} A + (N-1)B & \text{for } n = 0 \\ A - B & \text{for } n \neq 0 \end{cases}. \tag{22}$$

  (*Hint*: Recall that $\sum_{j=0}^{N-1} z_n^j = 0$ for all $n \neq 0$.)

(4) Fisher information for independent neurons.

Show that, in the uncorrelated case,

$$I_F^{\text{independent}} = \frac{1}{A} \sum_{j=0}^{N-1} \left(\frac{\partial m_j}{\partial s} \right)^2. \tag{23}$$

For a general (non-Gaussian) probability distribution, does the Fisher information also scale with population size, $I_F \sim \mathcal{O}(N)$? Why?

(5) Fisher information for neurons with uniform correlations.

Show that, in the uniform case,

$$I_F^{\text{uniform}} = \frac{I_F^{\text{independent}}}{1-c} - \frac{c}{A(1-c)[1+(N-1)c]} \left(\sum_{j=0}^{N-1} \frac{\partial m_j}{\partial s} \right)^2, \tag{24}$$

where we have defined the normalized correlation as

$$c \equiv \frac{B}{A}. \tag{25}$$

(*Hint*: Note that $M_0 = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \frac{\partial m_j}{\partial s}$.)

Thus, we have shown that in the limit of a large population,

$$I_F^{\text{uniform}} = \frac{I_F^{\text{independent}}}{1-c}, \tag{26}$$

i. e., uniform correlations enhance the Fisher information by a factor $(1-c)^{-1} \geq 1$.

(6) Fisher information for neurons with spatially decaying correlations, in the limit of a large population.

We assume actual *population* coding: for any given value of s a significant fraction of the population responds. In other words, we assume that individual tuning curves, $m_j(s)$, cover a finite fraction of the stimulus space. From this, it follows that only the first few M_ns survive and are of $\mathcal{O}\left(\sqrt{N}\right)$ (since they result from a sum of N terms normalized by $\sqrt{N}$); beyond some value of n of $\mathcal{O}(1)$, M_ns are negligible. The sum in Eq. (15) reduces to a sum over the first few terms only and, as a result, the scaling of I_F^{decaying} is dictated by the scaling of individual M_ns ($\sim \mathcal{O}\left(\sqrt{N}\right)$) and λ_ns for n small. Here, we elucidate the scaling of eigenvalues λ_ns for N large.

When we scale the population size, it is important that the N-dependence of the various terms make sense. In particular, when adding neurons we don't want to be increasing the size of the stimulus space; instead, we want to increase the density of neurons within a fixed stimulus space. Thus, we have to scale the distance between adjacent tuning curves, Δ, with N. Mathematically, we want

$$\Delta \to \frac{\Delta}{N}, \tag{27}$$

i. e.,

$$\rho = \rho_0^{1/N} \qquad \text{with} \qquad \rho_0 = \exp\left(-\frac{\Delta}{L}\right). \tag{28}$$

Show that, in the spatially decaying case, the correlation-dependent (*i. e.*, B-dependent) part of λ_n scales like N and calculate the prefactor. (*Hint*: Rewrite Eq. (21) in terms of a variable $u = 1/N$ and expand for small u. In order to obtain the dominant ($\mathcal{O}(1/u)$, equivalently $\mathcal{O}(N)$) term, you need only expand the numerator up to $\mathcal{O}(u)$ and the denominator up to $\mathcal{O}(u^2)$.)

Given this result, how does the Fisher information in the case of spatially decaying correlations, I_F^{decaying}, scale with population size, N?

(7) The most (only) important part.

Discuss your results. In particular, contrast the behaviors and relative magnitudes of $I_F^{\text{independent}}$, I_F^{uniform}, and I_F^{decaying}, and provide the intuitive pictures that you can think up.

Rava Azeredo da Silveira, Woods Hole, August 2009

Problem Set Solution

Simple Stuff in Information Theory

PROBLEM 1: Entropy of a Spike Train vs. Entropy of a Spike Count
(a) rT spikes are fired, on average, in the whole spike train. The number of possible arrangements of rT spikes in the T/τ bins is

$$\binom{T/\tau}{rT} = \frac{(T/\tau)!}{(rT)!\,(T/\tau - rT)!} \tag{1}$$

All the spike trains with rT spikes are equiprobable.
(b) The entropy is given by

$$\begin{aligned} H_{\text{train}} &= \ln\left(\#\text{ arrangements of } rT \text{ spikes in } T/\tau \text{ bins}\right) && (2)\\ &= \ln\left(\binom{T/\tau}{rT}\right) && (3)\\ &= \ln\left(\frac{(T/\tau)!}{(rT)!\,(T/\tau - rT)!}\right). && (4) \end{aligned}$$

According to Stirling's approximation, we have

$$\begin{aligned} H_{\text{train}} &\approx \frac{1}{\ln(2)} \ln\left(\frac{\left(\frac{T}{\tau e}\right)^{T/\tau}}{\left(\frac{rT}{e}\right)^{rT}\left(\frac{T/\tau - rT}{e}\right)^{T/\tau - rT}}\right) && (5)\\ &= \frac{1}{\ln(2)}\frac{T}{\tau}\ln\left(\frac{1/\tau}{r^{r\tau}\,(1/\tau - r)^{1-r\tau}}\right) && (6)\\ &= \frac{1}{\ln(2)}\frac{T}{\tau}\ln\left(\frac{1}{(r\tau)^{r\tau}\,(1-r\tau)^{1-r\tau}}\right) && (7)\\ &= -\frac{1}{\ln(2)}\frac{T}{\tau}\left[r\tau\ln(r\tau) + (1-r\tau)\ln(1-r\tau)\right]. && (8) \end{aligned}$$

If bins are independent, the entropy is the sum (hence the T/τ prefactor) of the entropy of each bin (hence the term in the square bracket).
(c) If $r\tau \ll 1$, we can use the approximation

$$\ln(1 - r\tau) \approx -r\tau. \tag{9}$$

Then

$$H_{\text{train}} \approx -\frac{1}{\ln(2)}\frac{T}{\tau}\left[r\tau\ln(r\tau)+(1-r\tau)(-r\tau)\right], \tag{10}$$

which reduces to

$$H_{\text{train}} \approx -\frac{1}{\ln(2)} rT\left[\ln(r\tau)-1\right] \tag{11}$$

when we ignore the $\mathcal{O}\left((r\tau)^2\right)$ on the right-hand-side. The entropy per spike is

$$H_{\text{per spike}} = \frac{H_{\text{train}}}{rT} \approx -\frac{1}{\ln(2)}\left[\ln(r\tau)-1\right], \tag{12}$$

which can be larger than one for small values of $r\tau$. This is because each spike occurs in a spike-free window of duration $\sim 1/r$ and, as there are $\sim 1/r\tau$ bins in this window,

$$H_{\text{per spike}} \sim \ln(\#\text{ states}) \sim \ln(1/r\tau). \tag{13}$$

(d) As we have seen in the lecture, the exponential distribution maximizes the entropy under the constraint of fixed mean:

$$P(k) = Z^{-1}e^{-\lambda k}, \tag{14}$$

where $k \geq 0$ is the spike count. Normalization imposes

$$1 = f(\lambda) \equiv \sum_{k=0}^{\infty} Z^{-1}e^{-\lambda k} = \frac{Z^{-1}}{1-e^{-\lambda}} \tag{15}$$

and the constraint on the mean imposes

$$rT = \langle k\rangle = -\partial_\lambda f(\lambda) = \frac{Z^{-1}e^{-\lambda}}{(1-e^{-\lambda})^2} = \frac{e^{-\lambda}}{1-e^{-\lambda}}. \tag{16}$$

Solving these two equations for Z and λ, we obtain

$$Z = 1+rT, \tag{17}$$

$$\lambda = \ln\left(1+\frac{1}{rT}\right). \tag{18}$$

Now, the entropy of this distribution is calculated as

$$H_{\text{count}} = -\frac{1}{\ln(2)}\langle\ln P(k)\rangle \tag{19}$$

$$= \frac{1}{\ln(2)}\left[\ln(Z)+\langle\lambda k\rangle\right] \tag{20}$$

$$= \frac{1}{\ln(2)}\left[\ln(Z)+\lambda rT\right] \tag{21}$$

$$= \frac{1}{\ln(2)}\left[\ln(1+rT)+rT\ln\left(1+\frac{1}{rT}\right)\right]. \tag{22}$$

(e) Typically, $H_{\text{train}} > H_{\text{count}}$ because the greater resolution at the level of individual spikes allows for a larger set of available states. In the limit with $rT \ll 1$, we can adopt the approximations

$$\ln(1+rT) \approx rT, \tag{23}$$

$$rT \ln\left(1+\frac{1}{rT}\right) \approx -rT \ln(rT), \tag{24}$$

so that the entropy becomes

$$H_{\text{count}} \approx -\frac{rT}{\ln(2)}\left[\ln(rT) - 1\right]. \tag{25}$$

The spike count entropy takes the form of a spike train entropy in this limit because the entire spike train consists (with high probability) of no more than a single spike.

PROBLEM 2: Equivalent Expressions of the Mutual Information

In the lecture, we defined the mutual information as

$$I = H(s) - H(s \mid r). \tag{26}$$

The conditional entropy can be decomposed as

$$\begin{aligned} H(s \mid r) &= -\sum_r P(r) \sum_s P(s \mid r) \ln\left(P(s \mid r)\right) && (27)\\ &= -\langle \ln\left(P(s \mid r)\right)\rangle && (28)\\ &= -\left\langle \ln\left(\frac{P(s,r)}{P(r)}\right)\right\rangle && (29)\\ &= -\langle \ln\left(P(s,r)\right)\rangle + \langle \ln\left(P(r)\right)\rangle && (30)\\ &= H(s,r) - H(r), && (31) \end{aligned}$$

such that

$$\begin{aligned} I &= H(s) - [H(s,r) - H(r)] && (32)\\ &= H(s) + H(r) - H(s,r). && (33) \end{aligned}$$

Since this expression is symmetric, it follows immediately that

$$I = H(r) - H(r \mid s). \tag{34}$$

Finally, we have also from above

$$\begin{aligned} I &= H(s) - H(s \mid r) && (35)\\ &= H(s) - \{[-\langle \ln\left(P(s,r)\right)\rangle] - [-\langle \ln\left(P(r)\right)\rangle]\} && (36)\\ &= H(s) - \left[\langle -\ln\left(P(r \mid s)\,P(s)\right)\rangle - \left\langle -\ln\left(\sum_s P(r \mid s)\,P(s)\right)\right\rangle\right] && (37) \end{aligned}$$

The first bracket is the information content of a r-state given a s-state, while the second bracket is the information content of a r-state regardless of the s-state. Thus, their difference is the 'ambiguity', *i. e.*, the information lost in the transformation of r-states into s-states.

PROBLEM 3: "Histogram Equalization"

(a) In the lecture, we showed that maximizing the entropy with no other constraint than the normalization of the probability distribution yields a uniform distribution.

(b) The distribution of s is given and the random variable m is related to s according to $m = m(s)$, so that $P(m)\,dm = P(s)\,ds$. We want to find the functional form $m(s)$ which yields a uniform distribution, $P(m) = \text{constant} \equiv c$. That is, we require

$$c dm = P(s)\,ds \tag{38}$$

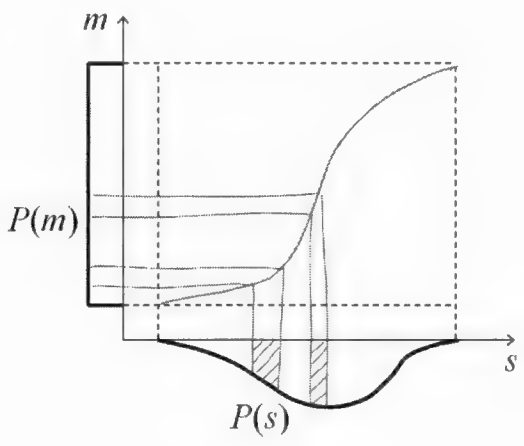
or

$$dm = \frac{1}{c} P(s)\,ds. \tag{39}$$

Integrating over s and keeping the normalization of $P(s)$ in mind, we obtain

$$m(s) = r_{\max} \int_{s_{\min}}^{s} d\tilde{s} dP(\tilde{s})\,. \tag{40}$$

(b) In the plot below, $m(s)$ is drawn in red. The two pairs of horizontal blue lines are equidistant. The shape of $m(s)$ is chosen such that the surfaces of the two shaded areas be equal.



PROBLEM 4: Maximizing the Mutual Information in a Linear Neuron

(a) Since r is a linear combination of Gaussian random variables, it is also itself a Gaussian random variable. with mean

$$\langle r \rangle = 0 \tag{41}$$

and variance

$$\begin{aligned}
\langle r^2 \rangle &= \left\langle \left(\sum_{i=1}^{N} w_i s_i + \delta \right)^2 \right\rangle && (42)\\
&= \sum_{i,j=1}^{N} w_i w_j \langle s_i s_j \rangle + \langle \delta^2 \rangle && (43)\\
&= \sum_{i,j=1}^{N} w_i w_j C_{ij} + \sigma^2 && (44)\\
&= w^T C w + \sigma^2. && (45)
\end{aligned}$$

Thus, the random variable r is distributed according to

$$P(r) = \frac{1}{\sqrt{2\pi \left(w^T C w + \sigma^2 \right)}} \exp\left(-\frac{1}{2} \frac{r^2}{w^T C w + \sigma^2} \right) \tag{46}$$

(b) The mutual information is calculated as

$$\begin{aligned}
I &= H(r) - H(r \mid s) && (47)\\
&= H(r) - H(\delta) && (48)\\
&= \frac{1}{2\ln(2)} \ln\left(2\pi e \left(w^T C w + \sigma^2 \right) \right) - \frac{1}{2\ln(2)} \ln\left(2\pi e \sigma^2 \right) && (49)\\
&= \frac{1}{2\ln(2)} \ln\left(1 + \frac{w^T C w}{\sigma^2} \right). && (50)
\end{aligned}$$

(c) The mutual information is maximal when the bilinear form $w^T C w$ is maximal. If we decompose w into the orthonormal eigenvectors of C,

$$w = \sum_{i=1}^{N} a_i e^{(i)}, \tag{51}$$

we find

$$\begin{aligned}
w^T C w &= \sum_{i=1}^{N} a_i a_j e^{(i)T} C e^{(j)} && (52)\\
&= \sum_{i=1}^{N} a_i a_j e^{(i)T} e^{(j)} \lambda_j && (53)\\
&= \sum_{i=1}^{N} a_i^2 \lambda_i. && (54)
\end{aligned}$$

If the norm of w is fixed, *i. e.*, if $\sum_{i=1}^{N} a_i^2$ is fixed, then $w^T C w$ is maximized by the choice $w^* = e^{(i^*)}$, where $e^{(i^*)}$ is the eigenvector associated to the largest eigenvalue.

(d) The choice $w = w^*$ corresponds to a neuron (*i. e.*, a response r) sensitive to the linear combination of stimuli along which these vary the most. That is, this choice exploits as much as possible the entropy in the stimuli.

Rava Azeredo da Silveira, Woods Hole, August 2009

PROBLEM SET SOLUTION

Fisher Information of Correlated Neurons

(1) General expression of the Fisher information.

Since the noise is additive, we can use the chain rule, $\frac{\partial}{\partial s} = -\sum_{i=0}^{N-1} \frac{\partial m_i}{\partial s} \frac{\partial}{\partial r_i}$, to obtain

$$
\begin{aligned}
I_F &= \left\langle -\frac{\partial^2 \ln P\left(r \mid s\right)}{\partial s^2} \right\rangle && (1)\\
&= -\left\langle \frac{\partial}{\partial s} \sum_{i=0}^{N-1} \frac{\partial m_i}{\partial s} \sum_{j=0}^{N-1} \left(C^{-1}\right)_{ij} \left[r_j - m_j\left(s\right)\right] \right\rangle && (2)\\
&= -\sum_{i=0}^{N-1} \frac{\partial^2 m_i}{\partial s^2} \sum_{j=0}^{N-1} \left(C^{-1}\right)_{ij} \left\langle r_j - m_j\left(s\right) \right\rangle && (3)\\
&\quad + \left\langle \sum_{i=0}^{N-1} \frac{\partial m_i}{\partial s} \sum_{j=0}^{N-1} \left(C^{-1}\right)_{ij} \frac{\partial m_j}{\partial s} \right\rangle && (4)\\
&= 0 + \sum_{i,j=0}^{N-1} \frac{\partial m_i}{\partial s} \left(C^{-1}\right)_{ij} \frac{\partial m_j}{\partial s}. && (5)
\end{aligned}
$$

(2) Expression of the Fisher information in terms of correlation eigenmodes.

Inverting the identity $V^T C V = \Lambda$, we find

$$\Lambda^{-1} = \left(V^T C V\right)^{-1} = V^{-1} C^{-1} \left(V^T\right)^{-1}. \quad (6)$$

We then obtain an expression for C^{-1} by 'sandwiching' this identity between V and V^T, as

$$V \Lambda^{-1} V^T = C^{-1}. \quad (7)$$

In the notation $m' = \frac{\partial m}{\partial s}$, the Fisher information becomes

$$
\begin{aligned}
I_F &= m'^T C^{-1} m' && (8)\\
&= m'^T V \Lambda^{-1} V^T m' && (9)\\
&= \begin{pmatrix} v^{(0)T}m' & \cdots & v^{(N-1)T}m' \end{pmatrix} \begin{pmatrix} \lambda_0^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_{N-1}^{-1} \end{pmatrix} \times && (10)\\
&\quad \times \begin{pmatrix} v^{(0)T}m' \\ \vdots \\ v^{(N-1)T}m' \end{pmatrix}, && (11)
\end{aligned}
$$

where the elements of the 'sandwiching' vectors are each dot products of the correlation matrix eigenvectors with the vector m'. Carrying out the matrix product, we obtain

$$
\begin{aligned}
I_F &= \begin{pmatrix} v^{(0)T}m' & \cdots & v^{(N-1)T}m' \end{pmatrix} \begin{pmatrix} \lambda_0^{-1} v^{(0)T}m' \\ \vdots \\ \lambda_{N-1}^{-1} v^{(N-1)T}m' \end{pmatrix} && (12)\\
&= \sum_{n=0}^{N-1} \frac{\left(v^{(n)T}m'\right)^2}{\lambda_n}. && (13)
\end{aligned}
$$

(3) Diagonalization of the covariance matrix.

Consider a circulant matrix,

$$
Q = \begin{pmatrix} q_0 & q_1 & \cdots & q_{N-2} & q_{N-1} \\ q_{N-1} & q_0 & \cdots & q_{N-3} & q_{N-2} \\ \vdots & & & & \vdots \\ q_1 & q_2 & \cdots & q_{N-1} & q_0 \end{pmatrix}, \qquad (14)
$$

and apply it to a vector

$$
W = \begin{pmatrix} 1 \\ w \\ \vdots \\ w^{N-2} \\ w^{N-1} \end{pmatrix}, \qquad (15)
$$

where w is a complex number:

$$
QW = \begin{pmatrix} \sum_{k=o}^{N-1} q_k w^k \\ q_{N-1} + \sum_{k=o}^{N-2} q_k w^{k+1} \\ \vdots \end{pmatrix}. \qquad (16)
$$

If $w^N = 1$, *i. e.*, if w is one of the N Nth root of unity, then the matrix identity becomes

$$QW = \left(\sum_{k=o}^{N-1} q_k w^k \right) \begin{pmatrix} 1 \\ w \\ \vdots \\ w^{N-2} \\ w^{N-1} \end{pmatrix}. \tag{17}$$

This demonstrates that W is an eigenvector of Q, with eigenvalue $\sum_{k=o}^{N-1} q_k w^k$. In order to normalize the eigenvector, we have to divide W by $\sqrt{N}$, since each of the complex numbers w^n is of unit norm.

In order to find the eigenvalues of the correlation matrix, we multiply the middle line (we assume N is odd) of the correlation matrix by the eigenvector, and divide the result by $z_n^{\frac{N-1}{2}}/\sqrt{N}$, to obtain

$$\lambda_n = \frac{C_0 z_n^{\frac{N-1}{2}}/\sqrt{N} + \sum_{j=1}^{\frac{N-1}{2}} C_j z_n^{\frac{N-1}{2}+j}/\sqrt{N} + \sum_{j=1}^{\frac{N-1}{2}} C_{-j} z_n^{\frac{N-1}{2}-j}/\sqrt{N}}{z_n^{\frac{N-1}{2}}/\sqrt{N}} \tag{18}$$

$$= C_0 + \sum_{j=1}^{\frac{N-1}{2}} C_j z_n^j + \sum_{j=1}^{\frac{N-1}{2}} C_{-j} z_n^{-j} \tag{19}$$

$$= C_0 + \sum_{j=1}^{\frac{N-1}{2}} C_j \left(z_n^j + z_n^{-j} \right). \tag{20}$$

- In the *uncorrelated case*, we know that all eigenvalues are identical and equal to A. It checks with the above formula for λ_n, since $C_0 = A$ and $C_j = 0$ for all $j \neq 0$.
- In the *spatially decaying case*, we obtain

$$\lambda_n = A + \sum_{j=1}^{\frac{N-1}{2}} B\rho^j \left(z_n^j + z_n^{-j} \right) \tag{21}$$

$$= A + 2B \,\mathrm{Re} \left(\sum_{j=1}^{\frac{N-1}{2}} \rho^j e^{i2\pi \frac{n}{N} j} \right) \tag{22}$$

$$= A + 2B \,\mathrm{Re} \left(\rho e^{i2\pi \frac{n}{N}} \frac{1 - \rho^{\frac{N-1}{2}} e^{i2\pi \frac{n}{N} \frac{N-1}{2}}}{1 - \rho e^{i2\pi \frac{n}{N}}} \right), \tag{23}$$

  which yields the expression given in the problem set sheet once the real part is calculated.

- In the *uniform case*,

$$\lambda_n = A + B \sum_{j=1}^{\frac{N-1}{2}} \left(z_n^j + z_n^{-j} \right) \tag{24}$$

$$= A - B + B \sum_{j=0}^{N-1} z_n^j \tag{25}$$

$$= \begin{cases} A - B + NB & \text{for } n = 0 \\ A - B + 0 & \text{for } n \neq 0 \end{cases} . \tag{26}$$

(4) Fisher information for independent neurons.
We had, in general,

$$I_F = \sum_{i,j=0}^{N-1} \frac{\partial m_i}{\partial s} \left(C^{-1} \right)_{ij} \frac{\partial m_j}{\partial s}. \tag{27}$$

Since in the uncorrelated case $\left(C^{-1} \right)_{ij} = \delta_{ij}/A$, this expression reduces to

$$I_F^{\text{independent}} = \frac{1}{A} \sum_{j=0}^{N-1} \left(\frac{\partial m_j}{\partial s} \right)^2 . \tag{28}$$

The linear growth of the Fisher information with population size, $I_F \sim \mathcal{O}(N)$, is valid generically in the case of independent neurons. This is because I_F is a linear function of $\ln(P)$ and, furthermore, $P = \prod_j P_j$ is a product over individual distributions if neurons are independent. As a result, I_F is the sum of N terms.

(5) Fisher information for neurons with uniform correlations.
We had obtained the general expression

$$I_F = \sum_{n=0}^{N-1} \frac{\left(v^{(n)T} m' \right)^2}{\lambda_n}. \tag{29}$$

Using the values of the eigenvalues in the uniform case, we can write

$$I_F^{\text{uniform}} = \frac{\left(v^{(0)T} m' \right)^2}{A + (N-1)B} + \sum_{n=1}^{N-1} \frac{\left(v^{(n)T} m' \right)^2}{A - B} \tag{30}$$

$$= \frac{\left(v^{(0)T} m' \right)^2}{A + (N-1)B} - \frac{\left(v^{(0)T} m' \right)^2}{A - B} + \sum_{n=0}^{N-1} \frac{\left(v^{(n)T} m' \right)^2}{A - B} \tag{31}$$

$$= \frac{A - B - A - (N-1)B}{(A-B)\left[A + (N-1)B\right]} \left(v^{(0)T} m' \right)^2 + \sum_{n=0}^{N-1} \frac{\left(v^{(n)T} m' \right)^2}{A - B} \tag{32}$$

$$= -\frac{NB}{(A-B)\left[A + (N-1)B\right]} \left(v^{(0)T} m' \right)^2 + \sum_{n=0}^{N-1} \frac{\left(v^{(n)T} m' \right)^2}{A - B} \tag{33}$$

But we know

$$v^{(0)T} = \frac{1}{\sqrt{N}}(1, ..., 1) \tag{34}$$

and

$$\sum_{n=0}^{N-1} \frac{\left(v^{(n)T} m'\right)^2}{A-B} = \frac{A}{A-B} I_F^{\text{independent}}, \tag{35}$$

so that

$$I_F^{\text{uniform}} = -\frac{B}{(A-B)\left[A+(N-1)B\right]} \left(\sum_{j=0}^{N-1} m'_j\right)^2 + \frac{A}{A-B} I_F^{\text{independent}}. \tag{36}$$

(6) Fisher information for neurons with spatially decaying correlations, in the limit of a large population.

In the spatially decaying case, we had found

$$\lambda_n = A + 2B\rho \frac{\cos(2\pi n/N) - \rho - (-1)^n \rho^{(N-1)/2}(1-\rho)\cos(\pi n/N)}{1 - 2\rho\cos(2\pi n/N) + \rho^2} \tag{37}$$

$$= A + 2B\rho_0^{1/N} \times \tag{38}$$

$$\times \frac{\cos(2\pi n/N) - \rho_0^{1/N} - (-1)^n \rho_0^{(1/N)(N-1)/2}\left(1-\rho_0^{1/N}\right)\cos(\pi n/N)}{1 - 2\rho_0^{1/N}\cos(2\pi n/N) + \rho_0^{2/N}} \tag{39}$$

$$= A + 2B\rho_0^u \frac{\cos(2\pi nu) - \rho_0^u - (-1)^n \rho_0^{1/2}\rho_0^{-u/2}(1-\rho_0^u)\cos(\pi nu)}{1 - 2\rho_0^u\cos(2\pi nu) + \rho_0^{2u}}, \tag{40}$$

where $u = 1/N$. Expanding the prefactor, the numerator, and the denominator for small u yields, respectively,

$$\rho_0^u = 1 + u\ln(\rho_0) + \mathcal{O}\left(u^2\right), \tag{41}$$

$$\text{numerator} = 1 - 1 - u\ln(\rho_0) \tag{42}$$

$$-(-1)^n \rho_0^{1/2}\left[1 - \frac{1}{2}u\ln(\rho_0)\right]\left[-u\ln(\rho_0)\right] + \mathcal{O}\left(u^2\right) \tag{43}$$

$$= -u\ln(\rho_0)\left[1 - (-1)^n\sqrt{\rho_0}\right] + \mathcal{O}\left(u^2\right), \tag{44}$$

$$\text{denominator} = 1 - 2\left\{1 + u\ln(\rho_0) + \frac{1}{2}\left[u\ln(\rho_0)\right]^2\right\}\left[1 - \frac{1}{2}(2\pi nu)^2\right] \tag{45}$$

$$+1 + 2u\ln(\rho_0) + \frac{1}{2}\left[2u\ln(\rho_0)\right]^2 + \mathcal{O}\left(u^3\right) \tag{46}$$

$$= 1 - 2 - 2u\ln(\rho_0) - \left[u\ln(\rho_0)\right]^2 + (2\pi nu)^2 \tag{47}$$

$$+1 + 2u\ln(\rho_0) + \frac{1}{2}\left[2u\ln(\rho_0)\right]^2 + \mathcal{O}\left(u^3\right) \tag{48}$$

$$= \left\{\left[\ln(\rho_0)\right]^2 + (2\pi n)^2\right\} u^2 + \mathcal{O}\left(u^3\right). \tag{49}$$

Combining all these, we get

$$\lambda_n = A + 2B\left[1 + u\ln(\rho_0)\right]\frac{-u\ln(\rho_0)\left[1-(-1)^n\sqrt{\rho_0}\right]}{\left\{\left[\ln(\rho_0)\right]^2 + (2\pi n)^2\right\}u^2} + \mathcal{O}\left(u^0\right) \quad (50)$$

$$= A\left\{1 - 2c\frac{\ln(\rho_0)\left[1-(-1)^n\sqrt{\rho_0}\right]}{\left\{\left[\ln(\rho_0)\right]^2 + (2\pi n)^2\right\}}\frac{1}{u} + \mathcal{O}\left(u^0\right)\right\}. \quad (51)$$

Remembering that $\rho_0 = e^{-\Delta/L}$ and $u = 1/N$, we can rewrite the $\mathcal{O}(N)$ contribution to the eigenvalue as

$$\lambda_n = 2Ac\frac{\Delta}{L}\frac{1-(-1)^n e^{-\Delta/2L}}{(\Delta/L)^2 + (2\pi n)^2} \times N + \mathcal{O}\left(N^0\right). \quad (52)$$

We can use this result to extract the scaling of the Fisher information as a function of N for large populations:

$$I_F^{\text{decaying}} = \sum_{n=0}^{N-1}\frac{\left(v^{(n)T}m'\right)^2}{\lambda_n} \quad (53)$$

$$\sim \frac{\left[\mathcal{O}\left(\sqrt{N}\right)\right]^2}{\mathcal{O}(N)} \quad (54)$$

$$\sim \mathcal{O}\left(N^0\right), \quad (55)$$

i. e., in large populations the variability of correlated modes ($\sim \mathcal{O}(N)$) is large enough to disturb the information contained in these modes (also $\sim \mathcal{O}(N)$).

(7) The most (only) important part.

See the intuitive pictures given in the lecture slides (or come up with more effective ones).

Applications of Spectral Methods in Functional Brain Imaging

David Kleinfeld[1] and Partha P. Mitra[2]

1- Department of Physics and Graduate Program in Neurosciences
University of California, La Jolla, CA
2- Cold Spring Harbor Laboratories, Cold Spring Harbor, NY

Prepared for:	Imaging in Neuroscience and Development, Rafael Yuste, series editor Book 1: Optical Imaging Techniques, Chapter 7, Rafael Yuste, editor Cold Spring Harbor Laboratory Press, NY
Narrative:	4000 words
Figures:	5 – all two columns in width with figures 3 and 5 in color
Figure captions:	800 words
References:	33 for 850 words
On-line material:	Two movies of material in figures 4A and 4B

Correspondence:	David Kleinfeld Department of Physics University of California at San Diego 9500 Gilman Drive La Jolla, CA 92093-0374 Office: 858-822-0342 Fax: 858-534-7697 Email: dk@physics.ucsd.edu

Abstract

Dynamic functional imaging experiments typically generate large, multivariate data sets that contain considerable spatial and temporal complexity. The goal of this chapter is to present signal processing techniques that allow the underlying spatio-temporal structure to be readily distilled and that further allow signal versus noise contributions to be separated.

Introduction

We present multivariate signal processing techniques that help reveal the spatio-temporal structure of optical imaging data, as well as allow signal versus noise contributions to be separated. These techniques typically assume that the underlying activity may be modeled as stationary stochastic processes over short analysis windows, *i.e.*, the statistics of the activity do not change during the analysis period. This requires selection of an appropriate temporal window for the analysis, which can be checked in a self-consistent manner.

We provide worked examples that serve to demonstrate the utility and implementation of spectral methods:

(1) Deduction of rhythmic components of the dilation and constriction of a cortical penetrating arteriole in rat to illustrate basic frequency-domain concepts.

(2) Deduction of synaptic connectivity between neurons in the leech swim network to emphasize the notions of spectral coherence and the associated confidence limits.

(3) The denoising of imaging data in the study of calcium waves in brain slice to introduce the concept of singular value decomposition (SVD) in the time domain and to illustrate the notion of space-time correlation in multi-site measurements.

(4) The delineation of wave phenomena in turtle visual cortex to illustrate spectrograms, along with the concept of SVD in the frequency domain to determine the dominant patterns of spatial coherence in a frequency localized manner.

Much of our exposition involves spectral analysis. Why work in the frequency domain?

- Many physiological phenomena have rhythmic components, ranging from electrical rhythms in the brain to visceral functions like breathing and heart beat. The time-series of these phenomena may appear very complicated, yet the representation in the frequency domain may be relatively simple and readily connected with underlying physiological processes.
- The calculation of confidence intervals requires that the number of degrees of freedom are known. Determining this number is complicated in the time domain, where all but white noise processes lead to correlation between neighboring data points. In contrast, counting the number of degrees of freedom is readily established in the frequency domain, as neighboring frequency bins are uncorrelated under stationarity assumptions.

Our emphasis is on the explanation and applications of signal processing methods and not on scientific questions *per se*.

Some relevant signal processing texts include Papoulis (1962), Ahmed and Rao (1975), and Percival and Walden (1993). The latter book includes sections on multitaper spectral analysis methods, developed originally by Thomson (1982) and used extensively in our analysis. Applications of modern signal processing methods to problems from neuroscience can be found in the book by Mitra and Bokil (2008) and in numerous reviews (Mitra and Pesaran 1998; Mitra et al. 1999; Pesaran et al. 2005; Kleinfeld 2008). Our notation follows that in Mitra and Bokil (2008).

Preliminaries

The process of data collection involves sampling a voltage or current so that the signals are represented as an ordered set of points, called a time series that are collected at a regular time interval. Let us denote the sampled time series as $\mathbf{V(t)}$, the sampling time interval as $\mathbf{\Delta t}$, and the length of sampling as $\mathbf{T}$. In spectral analysis, one re-expresses this time series in the frequency domain by decomposing $\mathbf{V(t)}$ into a weighted sum of sinusoids. We must first understand the range of frequencies that may be represented in the data.

The lowest resolvable frequency interval is given by the inverse of the length of the analysis window and is denoted the Raleigh frequency, $\mathbf{\Delta f_{Raleigh} = 1/T}$. In multitaper spectral analysis the

resolution bandwidth is typically denoted as $2\Delta\mathbf{f}$, where $\Delta\mathbf{f}$ is an adjustable parameter. The resolution bandwidth $2\Delta\mathbf{f}$ is also parameterized by the dimensionless product, **p**, of the half bandwidth $\Delta\mathbf{f}$ and the length of the window **T**, *i.e.*,

(1) $$\Delta f\, T = p$$

with $\mathbf{p} \geq 1$.

Sampling a continuous signal at discrete time intervals will in general lead to a loss of signal. However, there is an important class of signals, the so called band-limited signals whose spectral transforms vanish outside of a frequency range whose highest frequency is denoted **B**. For this case the signal can be perfectly reconstructed from discrete samples at uniform intervals $\Delta\mathbf{t}$, as long as the sampling interval is smaller than $\mathbf{1/(2B)}$. An alternative way of representing this criterion is to define the so called Nyquist frequency, $\mathbf{f_{Nyquist} = 1/(2\Delta t)}$. Then the criterion for perfect reconstruction of the original band-limited signal from sampled data becomes $\mathbf{f_{Nyquist} < B}$.

Neural signals are not naturally band-limited, although physiological mechanisms such as the membrane time constant of neurons provide natural cut-off frequencies. It is customary to lowpass filter the original analog signal so that it becomes effectively band-limited and the Nyquist criterion may be applied. Typically, $\mathbf{f_{Nyquist}}$ is chosen to be significantly larger than **B**, which is called "oversampling". However, if $\Delta\mathbf{t}$ is larger than $\mathbf{1/(2B)}$, the sampling is not sufficiently rapid and signals at frequencies higher than **B** are reflected back, or "aliased", into the sampled interval that ranges from 0 to $\mathbf{f_{Nyquist}}$. For example, a signal at 1.4-times $\mathbf{f_{Nyquist}}$ is aliased to appear at 0.6-times $\mathbf{f_{Nyquist}}$. The experimentally imposed lowpass filters to prevent aliasing are often called "antialiasing filters". Such filtering is not always possible and aliasing cannot always be avoided.

The discrete Fourier transform of a data segment is defined by

(2) $$\tilde{\mathbf{V}}(\mathbf{f}) = \frac{1}{\sqrt{\mathbf{T}}} \sum_{t=\Delta t}^{T} \Delta t\, e^{-i 2\pi f t} \mathbf{V}(\mathbf{t}).$$

Qualitatively, the time series **V(t)** is projected against all possible sinusoids, indexed by frequency **f**, to form a set of weights $\tilde{\mathbf{V}}(\mathbf{f})$. An immediate complication is that the finite extent of our data is equivalent to multiplying an infinite data series with a square pulse of width **T**. The

effect of such a window on the Fourier transform is to produce oscillations for each estimate of $\tilde{\mathbf{V}}(\mathbf{f})$ that extend into neighboring frequency bands (Figs. 1A and B). This is known as leakage, and is minimized by multiplying the time series with a function of time, denoted a taper, that smooths the sharp edges of the pulse. A half sine offers improvement and a Slepian taper is optimal; the cost of reduced leakage is decreased spectral resolution through an increase in the resolution bandwidth (Fig. 1B). There are considerable advantages to computing a family of independent estimates, $\tilde{\mathbf{V}}^{(\mathbf{k})}(\mathbf{f})$, rather than a single estimate, and in weighting the data with a set of orthonormal Slepian tapers to form this family (Fig. 1C). We denote each taper in the set by $\mathbf{w^{(k)}(t)}$ and compute the estimates

(3)
$$\tilde{\mathbf{V}}^{(\mathbf{k})}(\mathbf{f}) = \frac{\mathbf{1}}{\sqrt{\mathbf{T}}} \sum_{\mathbf{t}=\Delta\mathbf{t}}^{\mathbf{T}} \Delta\mathbf{t}\, \mathbf{e}^{-\mathbf{i}\,2\pi\mathbf{f}\mathbf{t}} \mathbf{w}^{(\mathbf{k})}(\mathbf{t}) \mathbf{V}(\mathbf{t}).$$

The maximum number of tapers, denoted **K**, that supports this minimization, and which is employed throughout our presentation, is

(4)
$$\mathbf{K = 2p - 1}.$$

The lower spectral resolution, or equivalently a larger resolution bandwidth, is offset by a greater number of spectral estimates $\tilde{\mathbf{V}}^{(\mathbf{k})}(\mathbf{f})$. The increase in number of independent estimates minimizes the distortion of the value in one frequency band by the value in a neighboring band and thus improves the statistical reliability of quantities that depend on the Fourier transform of the original signal.

Numerical processing of sampled data requires that we work in dimensionless units. We normalize time by the sample time, $\Delta\mathbf{t}$, so that the number of data points in the time series is given by $\mathbf{N \equiv T/\Delta t}$. We further normalize the resolution bandwidth by the sample time, $\Delta\mathbf{t}$, and define the unitless half-bandwidth $\mathbf{W \equiv \Delta t\, \Delta f}$. Then the time-frequency product $\mathbf{T\Delta f = p}$ is transformed to

(5)
$$\mathbf{NW = p}.$$

Given sampled data,

(6)
$$\tilde{\mathbf{V}}^{(\mathbf{k})}(\mathbf{f}) = \frac{\mathbf{1}}{\sqrt{\mathbf{N}}} \sum_{\mathbf{t}=\mathbf{1}}^{\mathbf{N}} \mathbf{e}^{-\mathbf{i}\,2\pi\mathbf{f}\mathbf{t}} \mathbf{w}_{\mathbf{t}}^{(\mathbf{k})}\, \mathbf{V}_{\mathbf{t}}$$

where time is now an index that runs from **1** to **N** in steps of **1** rather than a discrete variable that runs from **0** to **T** in steps of Δt, while frequency runs from **-½** to **+½** in steps of **1/N** (Table 1).

Implementation of the algorithms can be in any programming environment, but the use of the MatLab™-based programming environment along with packaged routines in Chronux (www.chronux.org) is particularly convenient.

Case one: Spectral power

As a means to introduce spectral estimation, we analyze the rhythms that give rise to motion of the wall of a penetrating arteriole that sources blood to cortex (Fig. 2A). These arterioles are gateways that transfer blood from the surface of cortex to the underlying microvasculature (Nishimura et al. 2007). Past work has established that isolated arterioles can generate myographic activity in the 0.1 Hz range (Osol and Halpern 1988), similar to that seen *in vivo* by noninvasive imaging techniques of brain tissue (Mayhew et al. 1996) and two-photon imaging of capillaries (Kleinfeld et al. 1998). Here we look at the relative contribution of vasomotion, as well as breathing and heart rate, to penetrating vessels.

The raw signal is the diameter of the vessel, denoted $\mathbf{D(t)}$ (Fig. 2B). The mean value is removed to give

$$\delta D_t = D_t - \frac{1}{N}\sum_{t=1}^{N} D_t \,. \tag{7}$$

Our goal is to understand the spectral content of this signal – with confidence limits! The Fourier transform of this signal, with respect to the k^{th} taper, is (Eq. 6)

$$\delta\tilde{D}^{(k)}(f) = \frac{1}{\sqrt{N}}\sum_{t=1}^{N} e^{-i2\pi f t} w_t^{(k)}\, \delta D_t \tag{8}$$

where, as noted previously (Eq. 3), $\mathbf{w^{(k)}(t)}$ is the k^{th} Slepian taper, whose length is also **T**. We compute the spectral power density, denoted $\tilde{S}(f)$, with units of distance^2/frequency, in terms of an average over the index of tapers, *i.e.*,

$$\tilde{S}(f) \equiv \frac{1}{K}\sum_{k=1}^{K}\left|\delta\tilde{D}^{(k)}(f)\right|^2 \tag{9}$$

where $\left|\delta\tilde{\mathbf{D}}^{(k)}(\mathbf{f})\right|^2 = \delta\tilde{\mathbf{D}}^{(k)}(\mathbf{f})\left[\delta\tilde{\mathbf{D}}^{(k)}(\mathbf{f})\right]^*$; we further average over all trials if appropriate. Note that the "$\mathbf{1}/\sqrt{\mathbf{N}}$" normalization satisfies Parseval's theorem, *i.e.*,

$$\sum_{f=0}^{f_{Nyquist}} \tilde{S}(f) = \frac{1}{N}\sum_{t=1}^{N} D_t^2 . \quad \textbf{(10)}$$

The spectrum in this example has strong features, yet has a trend to decrease as roughly $1/f^2$ that tends to obscure the peaks (insert in Fig. 2C). We remove the trend by computing the spectrum of the temporal derivative of δD_t,

$$\delta D'_t \equiv \frac{\delta D_{t+1} - \delta D_t}{\Delta t}, \quad \textbf{(11)}$$

as a means to flatten or "prewhiten" the spectrum. We now observe a multitude of peaks on a relatively flat background (Fig. 2C). A broad peak is centered at 0.2 Hz and corresponds to vasomotion. A sharper peak near 1 Hz corresponds to breathing; the nonsinusoidal shape of variations in diameter caused by breathing leads to the presence of second and third harmonics. Finally, a sharp peak at 7 Hz corresponds to heart rate and also includes a harmonic. No additional peaks are observed beyond the second harmonic of breathing. Note that the resolution half-bandwidth was chosen to be $\Delta f = 0.03$ Hz ($p = 16$), which is narrower than the low frequency band.

The next issue is the calculation of confidence intervals so that the uncertainty in the power at each peak may be established and the statistical significance of each peak may be assessed. Confidence limits may be estimated analytically for various asymptotic limits. However, the confidence intervals may also be estimated directly by a jack-knife, where we compute the standard error in terms of "delete-one" means (Thomson and Chave 1991). In this procedure, we exploit the multiple estimates of the spectral power density and calculate **K** different mean spectra in which one term is left out, *i.e.*,

$$\tilde{S}^{(n)}(f) \equiv \frac{1}{K-1}\sum_{\substack{k=1 \\ k\neq n}}^{K}\left|\delta\tilde{D}^{(k)}(f)\right|^2 . \quad \textbf{(12)}$$

Estimating the standard error of the spectral power density requires an extra step since spectral amplitudes are defined on the interval [0, ∞) while Gaussian variables exist on the full interval

(-∞, ∞). Taking the logarithm leads to variables defined over the full interval; thus we transform the delete-one estimates, $\tilde{\mathbf{S}}^{(n)}(\mathbf{f})$, according to

(13) $$\mathbf{g}\left\{\tilde{\mathbf{S}}^{(n)}(\mathbf{f})\right\} \equiv \mathit{ln}\left\{\tilde{\mathbf{S}}^{(n)}(\mathbf{f})\right\}$$

or

(14) $$\tilde{\mathbf{S}}^{(n)}(\mathbf{f}) = \mathbf{e}^{\mathbf{g}\left\{\tilde{\mathbf{S}}^{(n)}(\mathbf{f})\right\}} .$$

The mean of the transformed variable is

(15) $$\tilde{\mu}(\mathbf{f}) \equiv \frac{\mathbf{1}}{\mathbf{K}} \sum_{\mathbf{k}=\mathbf{1}}^{\mathbf{K}} \mathbf{g}\left\{\tilde{\mathbf{S}}^{(n)}(\mathbf{f})\right\}$$

and the standard error is

(16) $$\tilde{\sigma}(\mathbf{f}) = \sqrt{\frac{\mathbf{K}-\mathbf{1}}{\mathbf{K}} \sum_{\mathbf{n}=\mathbf{1}}^{\mathbf{K}} \left[\mathbf{g}\left\{\tilde{\mathbf{S}}^{(n)}(\mathbf{f})\right\} - \mu(\mathbf{f})\right]^2} .$$

The "95 %" confidence limit for the transformed spectral density is given by $2\tilde{\sigma}(\mathbf{f})$, so that one visualizes $\tilde{\mathbf{S}}(\mathbf{f})$ by plotting the mean value of $\tilde{\mathbf{S}}(\mathbf{f})$, *i.e.*, $\mathbf{e}^{\tilde{\mu}(\mathbf{f})}$, along with the lower and upper bounds, *i.e.*, $\mathbf{e}^{\tilde{\mu}(\mathbf{f})-2\tilde{\sigma}(\mathbf{f})}$ and $\mathbf{e}^{\tilde{\mu}(\mathbf{f})+2\tilde{\sigma}(\mathbf{f})}$, respectively. The confidence bands are symmetric about the mean when spectral power is plotted on a logarithmic scale (upper trace in Fig. 2C) rather than on a linear scale (lower trace in Fig. 2C).

Case two: Coherence between two signals

As a means to introduce coherence, a measure of the tracking of one rhythmic signal by another, we consider the use of optical imaging to determine potential pair-wise connections between neurons (Cacciatore et al. 1999). We focus on imaging data taken from the ventral surface of a leech ganglion and seek to identify cells in the ganglion that receive monosynaptic input from neuron Tr2 in the head (Fig. 3A). This cell functions as a toggle for the swim rhythm in these animals. Rather than serially impale each of the roughly 400 cells in the ganglion and look for post-synaptic currents induced by driving Tr2, a parallel strategy was adopted (Taylor et al.

2003). The cells in the ganglion were stained with a voltage-sensitive dye (Fig. 3B), which transforms changes in membrane potential into changes in the intensity of fluorescent light. The emitted light from all cells is detected with a charge-coupled device (Fig. 3B), from which time series for the change in fluorescence are calculated for each neuron in the field (Fig. 3C). Presynaptic cell Tr2 was stimulated with a periodic signal, at frequency $\mathbf{f_{Drive}}$, with the assumption that candidate post-synaptic followers of Tr2 would fire with the same periodicity (Fig. 3D). The phase of the coherence relative to the drive depends on the sign of the synapse, propagation delays, and filtering by postsynaptic processes.

The coherence between the response of each cell and the drive, a complex function denoted $\tilde{\mathbf{C}}_i(\mathbf{f})$, was calculated over the time period of the stimulus. We denote the measured time series of the optical signals as $\mathbf{O_i(t)}$ and the reference drive signal as $\mathbf{R(t)}$. The spectral coherence is defined as

$$\tilde{\mathbf{C}}_i(\mathbf{f}) = \frac{\frac{1}{\mathbf{K}}\sum_{\mathbf{k}=1}^{\mathbf{K}} \tilde{\mathbf{O}}_i^{(\mathbf{k})}(\mathbf{f})\left[\tilde{\mathbf{R}}_i^{(\mathbf{k})}(\mathbf{f})\right]^*}{\sqrt{\left(\frac{1}{\mathbf{K}}\sum_{\mathbf{k}=1}^{\mathbf{K}}\left|\tilde{\mathbf{O}}_i^{(\mathbf{k})}(\mathbf{f})\right|^2\right)\left(\frac{1}{\mathbf{K}}\sum_{\mathbf{k}=1}^{\mathbf{K}}\left|\tilde{\mathbf{R}}_i^{(\mathbf{k})}(\mathbf{f})\right|^2\right)}}. \tag{17}$$

To calculate the standard errors for the coherence estimates, we again use the jack-knife (Thomson and Chave 1991) and compute "delete-one" averages of coherence, denoted $\tilde{\mathbf{C}}_i^{(\mathbf{n})}(\mathbf{f})$ where $\mathbf{n}$ is the index of the deleted taper, *i.e.*,

$$\tilde{\mathbf{C}}_i^{(\mathbf{n})}(\mathbf{f}) = \frac{\frac{1}{\mathbf{K}-1}\sum_{\substack{\mathbf{k}=1\\ \mathbf{k}\neq\mathbf{n}}}^{\mathbf{K}} \tilde{\mathbf{O}}_i^{(\mathbf{k})}(\mathbf{f})\left[\tilde{\mathbf{R}}_i^{(\mathbf{k})}(\mathbf{f})\right]^*}{\sqrt{\left(\frac{1}{\mathbf{K}-1}\sum_{\substack{\mathbf{k}=1\\ \mathbf{k}\neq\mathbf{n}}}^{\mathbf{K}}\left|\tilde{\mathbf{O}}_i^{(\mathbf{k})}(\mathbf{f})\right|^2\right)\left(\frac{1}{\mathbf{K}-1}\sum_{\substack{\mathbf{k}=1\\ \mathbf{k}\neq\mathbf{n}}}^{\mathbf{K}}\left|\tilde{\mathbf{R}}_i^{(\mathbf{k})}(\mathbf{f})\right|^2\right)}}. \tag{18}$$

Estimating the standard error of the magnitude of $\tilde{\mathbf{C}}_i(\mathbf{f})$, as with the case for the spectral power, requires an extra step since $\left|\tilde{\mathbf{C}}_i(\mathbf{f})\right|$ is defined on the interval [0, 1] while Gaussian variables exist on $(-\infty, \infty)$. The mean value of the magnitude of the coherence for each postsynaptic cell, *i.e.*, $\left|\tilde{\mathbf{C}}_i(\mathbf{f})\right|$, and the "delete-one" estimates, $\left|\tilde{\mathbf{C}}_i^{(\mathbf{n})}(\mathbf{f})\right|$, are replaced with the transformed values

(19) $$\mathbf{g}\left\{\left|\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right|\right\} = ln\left(\frac{\left|\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right|^2}{1 - \left|\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right|^2}\right)$$

or

(20) $$\left|\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right| = \frac{\mathbf{1}}{\sqrt{\mathbf{1} + e^{-\mathbf{g}\left\{\left|\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right|\right\}}}}.$$

The means of the transformed variables are

(21) $$\tilde{\mu}_{i;\,\mathrm{Mag}}(\mathbf{f}) = \frac{\mathbf{1}}{\mathbf{K}}\sum_{n=1}^{K}\mathbf{g}\left\{\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right\}.$$

and their standard errors are

(22) $$\tilde{\sigma}_{i;\,\mathrm{Mag}}(\mathbf{f}) = \sqrt{\frac{\mathbf{K}-\mathbf{1}}{\mathbf{K}}\sum_{n=1}^{K}\left|\mathbf{g}\left\{\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right\} - \tilde{\mu}_{i;\,\mathrm{Mag}}(\mathbf{f})\right|^2}.$$

The "95 %" confidence interval for $\left|\tilde{\mathbf{C}}_i(\mathbf{f})\right|$ corresponds to values within the interval $\left[\sqrt[-1]{\mathbf{1}+e^{-(\tilde{\mu}_{i;\mathrm{Mag}} - 2\tilde{\sigma}_{i;\mathrm{Mag}})}}, \sqrt[-1]{\mathbf{1}+e^{-(\tilde{\mu}_{i;\mathrm{Mag}} + 2\tilde{\sigma}_{i;\mathrm{Mag}})}}\right]$. For completeness, an alternate transformation for computing the variance is $\mathbf{g}\left\{\tilde{\mathbf{C}}_i(\mathbf{f})\right\} = \mathbf{tanh}^{-1}\left\{\tilde{\mathbf{C}}_i(\mathbf{f})\right\}$.

We now consider an estimate of the standard deviation of the phase of $\tilde{\mathbf{C}}_i(\mathbf{f})$. Conceptually, the idea to compute the variation in the relative directions of the "delete-one" unit vectors, *i.e.*, $\tilde{\mathbf{C}}_i(\mathbf{f})/\left|\tilde{\mathbf{C}}_i(\mathbf{f})\right|$. The standard error is computed as

(23) $$\tilde{\sigma}_{i;\,\mathrm{Phase}}(\mathbf{f}) = \sqrt{2\,\frac{\mathbf{K}-\mathbf{1}}{\mathbf{K}}\left(\mathbf{K} - \left|\sum_{n=1}^{K}\frac{\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})}{\left|\tilde{\mathbf{C}}_i^{(n)}(\mathbf{f})\right|}\right|\right)}.$$

Our interest is in the values of $\tilde{\mathbf{C}}_i(\mathbf{f})$ for $\mathbf{f} = \mathbf{f}_{\mathbf{Drive}}$ and the confidence limits for this value. We choose the resolution bandwidth so that the estimate of $\left|\tilde{\mathbf{C}}_i(\mathbf{f}_{\mathbf{Drive}})\right|$ is kept separate from that

of the harmonic $\left|\tilde{C}_i\left(2f_{Drive}\right)\right|$; the choice $\Delta f = 0.4\, f_{Drive}$ works well. We graph the magnitude and phase of $\tilde{C}_i\left(f_{Drive}\right)$ for all neurons, along with the confidence intervals, on a polar plot (Fig. 3E).

Finally, we consider whether the coherence of a given cell at f_{Drive} is significantly greater than zero, *i.e.*, larger than one would expect by chance from a signal with no coherence, as a means to select candidate post-synaptic targets of Tr2. We compared the estimate for each value of $\tilde{C}_i\left(f_{Drive}\right)$ to the null distribution for the magnitude of the coherence, which exceeds

(24) $$\left|\tilde{C}_i\left(f_{Drive}\right)\right| = \sqrt{1 - \alpha^{1/(K-1)}}$$

only in a fraction α of the trials (Hannan 1970; Jarvis and Mitra 2001). We used $\alpha = 0.001$ in our experiments to avoid false positives. We also calculated the multiple comparisons α level for each trial, given by

(25) $$\alpha_{multi} = 1 - (1 - \alpha)^N,$$

where N is the number of cells in the functional image, and verified that it did not exceed $\alpha_{multi} = 0.05$ on any trial (Fig. 3E).

The result of the above procedure was the discovery of three postsynaptic targets of cell Tr2, two of which were functionally unidentified neurons (Taylor et al. 2003).

Case three: Space-time singular valued decomposition and denoising

A common issue in the analysis of optical imaging data is the need to remove "fast" noise, that is, fluctuations in intensity that occur on a pixel-by-pixel and frame-by-frame basis. The idea is that the imaging data contains features that are highly correlated in space, such as underlying cell bodies, processes, *etc.*, and highly correlated in time, such as long-lasting responses. The imaging data may thus be viewed as a space-time matrix of random numbers, *i.e.*, the fast noise, with added correlated structure. The goal is to separate the fast, uncorrelated noise from the raw data so that a compressed image file with only the correlated signals remain (Figs. 4A and B show single frames; the complete movies are available online). With this model in mind, we focus on the case of intracellular Ca^{2+} oscillations in an organotypic culture of rat cortex, which contains both neurons and glia. All cells were loaded with a fluorescence-based calcium

indicator, and spontaneous activity in the preparation was imaged with a fast-framing ($\Delta t = 2$ ms), low resolution (100 by 100 pixels) confocal microscope (Fig. 4A).

Imaging data is in the form of a 3-dimensional array of intensities, denoted $\mathbf{V(x, y, t)}$. We consider expressing the spatial location in terms of a pixel index, so that each $\mathbf{(x, y) \rightarrow s}$ and the data is now in the form of a space-time matrix, *i.e.*, $\mathbf{V(s, t)}$. This matrix may be decomposed into the outer product of functions of pixel index with functions of time. Specifically,

$$\mathbf{V(s, t)} = \sum_{n=1}^{rank\{V\}} \lambda_n \, \mathbf{F_n(s)} \, \mathbf{G_n(t)} \tag{26}$$

where the rank of $\mathbf{V(s, t)}$ is the smaller of the pixel or time dimensions. For example data of Figure 4A, there are $\mathbf{N_t} = 1{,}200$ frames, or time points, and $\mathbf{N_s} = 10{,}000$ pixels, so that $\mathbf{rank\{V(s, t)\} = N_t}$. The above decomposition is referred to as a singular value decomposition (Golub and Kahan 1965). The temporal functions satisfy an eigenvalue equation, *i.e.*,

$$\sum_{t'=1}^{N_t} \mathbf{G_n(t')} \left[\sum_{s=1}^{N_s} \mathbf{V(s, t) V(s, t')} \right] = \lambda_n^2 \mathbf{G_n(t)} \tag{27}$$

where the functions $\mathbf{F_n(s)}$ and $\mathbf{G_n(t)}$ are orthonormal. The spatial function that accompanies each temporal function is found by inverting the defining equation, so that

$$\mathbf{F_n(s)} = \frac{1}{\lambda_n} \sum_{t=1}^{N_t} \mathbf{V(s, t) \, G_n(t)} \; . \tag{28}$$

For completeness, note that this is equivalent to determining the principal components of responses recorded from a single location across multiple trials, as opposed to multiple locations in a single trial, where **s** labels the trial rather than the location.

When this decomposition is applied to the Ca^{2+} imaging data (Fig. 4A), we see that the eigenvalue spectrum has large values for the low-order modes and then rapidly falls to a smoothly decreasing function of index (Fig. 4B); theoretical expressions for the baseline distribution have been derived (Sengupta and Mitra 1999). The spatial and temporal modes show defined structure for the first ~20 modes; beyond this the spatial modes appear increasingly "grainy" and the temporal modes appear as fast noise (Figs. 4D and E).

The utility of this decomposition is that only the lower order modes carry information. Thus we can reconstruct the data matrix from only these modes and remove the "fast" noise, *i.e.*,

$$\mathbf{V}^{\text{reconstructed}}(\mathbf{s}, \mathbf{t}) = \sum_{n=1}^{\text{largest significant mode}} \lambda_n \, \mathbf{F}_n(\mathbf{s}) \, \mathbf{G}_n(\mathbf{t}) . \tag{29}$$

Compared to smoothing techniques, the truncated reconstruction respects all correlated features in the data and thus, *e.g.*, does not remove sharp edges. Reconstruction of the intracellular Ca^{2+} oscillation data highlights the correlated activity by removing fast, grainy variability (Fig. 4B).

Case four: Spectrograms and space-frequency singular value decomposition

The final example concerns the characterization of coherent spatio-temporal dynamics, such as waves of activity. We return to the use of voltage-sensitive dyes, this time to image the electrical dynamics of turtle visual cortex in response to a looming stimulus. Early work had shown that a looming stimulus led to the onset of ~ 20 Hz oscillations, the g-band for turtle, in visual cortex (Prechtl and Bullock 1994;1995). The limited range of cortical connections led to the hypothesis that this oscillation might be part of wave motion. We investigated this issue by direct electrical measurements throughout the depth of cortex at selected sites (Prechtl et al. 2000) and, of relevance for the present discussion, by imaging the spatial patterns from cortex using voltage sensitive dyes as the contrast agent (Prechtl et al. 1997).

The electrical activity is expected to evolve between pre-stimulus versus post-stimulus epochs, and possibly over an extended period of stimulation. Thus the spectral power is not stationary over long periods of time and we must consider a running measure of the spectral power density, denoted the spectrogram, that is a function of both frequency and time. We choose a restricted interval of time, denoted $\mathbf{T_{window}}$ with $\mathbf{N_{window}}$ data points, compute the Fourier transforms $\tilde{\mathbf{V}}^{(k)}(\mathbf{f}; \mathbf{t}_0)$ and spectrum $\tilde{\mathbf{S}}(\mathbf{f}; \mathbf{t}_0)$ over that interval, where $\mathbf{t_0}$ indexes the time at the middle of the epoch, and then step forward in time and recalculate the transforms and spectrum. Thus

$$\tilde{\mathbf{V}}^{(k)}(\mathbf{f}; \mathbf{t}_0) = \frac{1}{\sqrt{\mathbf{N}}} \sum_{t = t_0 - \frac{1}{2}N_{\text{window}}}^{t_0 + \frac{1}{2}N_{\text{window}} - 1} \mathbf{e}^{-i 2\pi f t} \mathbf{w}_t^{(k)} \, \mathbf{V}_t \tag{30}$$

and

$$\tilde{S}(f; t_0) \equiv \frac{1}{K} \sum_{k=1}^{K} \left| \tilde{V}^{(k)}(f; t_0) \right|^2 . \tag{31}$$

The resolution half-bandwidth is now p/N_{window} and, as a practical matter, the index is shifted in increments no larger than $N_{window}/2$. For the case of the summed optical signal from turtle cortex, we observe low frequency activity prior to stimulation and multiple bands of high frequency oscillations upon stimulation (Fig. 5A). A particularly pronounced band occurs near 18 Hz; this is clearly seen in a line plot of the spectral power density for the one second epoch centered in the middle of the stimulation period (insert in Fig. 5B).

The image data, even after denoising (Eq. 29) and broad-band filtering, appears complex (Fig. 5B), with regions of net depolarization sweep across cortex, but no simple pattern emerges. One possibility is that cortex supports multiple dynamic processes, each with a unique center frequency, that may be decomposed by a singular value decomposition in the frequency domain. In this method, proposed by Mann and Park (1994), the space-time data $V(s, t)$ is first projected into a local temporal frequency domain by transforming with respect to a set of tapers, *i.e.*,

$$\tilde{V}^{(k)}(s, f) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} e^{-i 2\pi f t} w_t^{(k)} V_t(s) . \tag{32}$$

The index k defines a local frequency index in the band $[f_0 - \Delta f/2,\ f_0 + \Delta f/2]$. For a fixed frequency, f_0, a SVD is performed on the complex matrix

$$\tilde{V}(s, k; f_0) \equiv \left[\tilde{V}^{(1)}(s, f_0), ..., \tilde{V}^{(K)}(s, f_0) \right] \tag{33}$$

to yield

$$\tilde{V}(s, k; f_0) = \sum_{n=1}^{\text{rank}\{\tilde{V}\}} \lambda_n \tilde{F}_n(s) \tilde{G}_n(k) \tag{34}$$

where the rank is invariably set by K. A measure of coherence is given by the ratio of the power of the leading mode to the total power (Fig. 5C), *i.e.*,

$$\hat{C}(f_0) = \frac{\lambda_1^2(f_0)}{\sum_{k=1}^{K} \lambda_k^2(f_0)} . \tag{35}$$

A completely coherent response leads to $\hat{C}(f_0) = 1$, while for a uniform random process $\hat{C}(f_0) = 1/K$. Where $\hat{C}(f_0)$ has a peak, it is useful to examine the largest spatial mode, $\tilde{F}_1(s)$. The magnitude of this complex image gives the spatial distribution of coherence that is centered at frequency f_0, while gradients in the phase of the image indicate the local direction of propagation.

For the example data (Fig. 5B), this analysis revealed linear waves as the dominant mode of electrical activity. Those with a temporal frequency centered at f_0 = 3 Hz are present with or without stimulation (Prechtl et al. 1997) (Fig. 5D) while those centered at f_0 = 8, 18, and 23 Hz are seen only with stimulation and propagate orthogonal to the wave at 3 Hz (Fig. 5E-G). It is of biological interest that the waves at f_0 = 3 Hz track the direction of thalamo-cortical input, while those at higher frequencies track a slight bias in axonal orientation (Cosans and Ulinski 1990) that was unappreciated in the original work (Prechtl et al. 1997).

Epilog

Our presentation covered a number of key applications of spectral methods to optical imaging data. The choice of topics is representative but by no means exhaustive. An additional application that is likely to be of utility is the fitting of line-spectra to signals with relatively pure periodic contributions, such as may occur from physiological rhythms, the response to a periodic stimulus, or from environmental contaminants like line power (Mitra et al. 1999; Pesaran et al. 2005). A second application of note is demodulation of a spatial image in response to periodic stimulation either at the fundamental drive frequency (Borst 1995; Kalatsky and Stryker 2003; Sornborger et al. 2005) or the second harmonic of the drive (Benucci et al. 2007). Demodulation also is valuable for delineating wave dynamics in systems with rhythmic activity (Kleinfeld et al. 1994; Prechtl et al. 1997). In general, spectral techniques are an essential tool for the statistical analysis of imaging data.

Acknowledgements

We thank Bijan Pesaran and David J. Thomson for many useful discussions, Andy Y. Shih for acquiring the unpublished data in figure 2, Joshua T. Vogelstein for acquiring the data in figure 4, and Adrienne L.

Fairhall and Karel Svoboda for comments on a preliminary version of the chapter. The material is derived from a presentation at the Society for Neuroscience short course on "Neural Signal Processing: Quantitative Analysis of Neural Activity" as well as presentations at the "Neuroinformatics" and "Methods in Computational Neuroscience" schools at the Marine Biology Laboratories and the "Imaging Structure and Function in the Nervous System" school at Cold Spring Harbor Laboratory. The development of spectral tools in the Chronux library was funded by the NIH (grant MH071744 to PPM) while the application of spectral methods to imaging data sets was also funded by the NIH (grants EB003832, MH085499, and NS059832 to DK and MH062528 to PPM).

References

Ahmed, N. and Rao, K.R. 1975. *Orthogonal transforms for Digital Signal Processing*. Springer Verlag, New York.

Benucci, A., Frazor, R.A., and Carandini, M. 2007. Standing waves and traveling waves distinguish two circuits in visual cortex. *Neuron* 55: 103-117.

Borst, A. 1995. Periodic current injection (PCI): A new method to image steady-state membrane potential of single neurons in situ using extracellular voltage-sensitive dyes. *Zeitschrift für Naturforschung* 50: 435-438.

Cacciatore, T.W., Brodfueher, P.D., Gonzalez, J.E., Jiang, T., Adams, S.R., Tsien, R.Y., Kristan Jr., W.B., and Kleinfeld, D. 1999. Identification of neural circuits by imaging coherent electrical activity with FRET-based dyes. *Neuron* 23: 449-459.

Cosans, C.E. and Ulinski, P.S. 1990. Spatial organization of axons in turtle visual cortex: Intralamellar and interlamellar projections. *Journal of Comparative Neurology* 296: 548-558.

Devor, A., Tian, P., Nishimura, N., Teng, I.C., Hillman, E.M., Narayanan, S.N., Ulbert, I., Boas, D.A., Kleinfeld, D., and Dale, A.M. 2007. Suppressed neuronal activity and concurrent arteriolar vasoconstriction may explain negative blood oxygenation level-dependent signaling. *Journal of Neuroscience* 27: 4452-4459.

Golub, G.H. and Kahan, W. 1965. *Calculating Singular Values and Pseudo-Inverse of a Matrix*. Society for Industrial and Applied Mathematics, Philadelphia.

Hannan, E.J. 1970. *Multiple Time Series*. Wiley, New York.

Jarvis, M.R. and Mitra, P.P. 2001. Sampling properties of the spectrum and coherency of sequences of action potentials. *Neural Computation* 13: 717-749.

Kalatsky, V.A. and Stryker, M.P. 2003. New paradigm for optical imaging: Temporally encoded maps of intrinsic signal. *Neuron* 38: 529-545.

Kleinfeld, D. 2008. Application of spectral methods to representative data sets in electrophysiology and functional neuroimaging. in *Syllabus for Society for Neuroscience Short Course III on "Neural Signal Processing: Quantitative Analysis of Neural Activity"* (ed. P. Mitra), pp. 21-34. Society for Neuroscience, Washington, DC.

Kleinfeld, D., Delaney, K.R., Fee, M.S., Flores, J.A., Tank, D.W., and Gelperin, A. 1994. Dynamics of propagating waves in the olfactory network of a terrestrial mollusk: An electrical and optical study. *Journal of Neurophysiology* 72: 1402-1419.

Kleinfeld, D., Mitra, P.P., Helmchen, F., and Denk, W. 1998. Fluctuations and stimulus-induced changes in blood flow observed in individual capillaries in layers 2 through 4 of rat neocortex. *Proceedings of the National Academy of Sciences USA* 95: 15741-15746.

Mann, M.E. and Park, J. 1994. Global-scale modes of surface temperature variability on interannual to centuries timescales. *Journal of Geophysical Research* 99: 25819-25833.

Mayhew, J.E.W., Askew, S., Zeng, Y., Porrill, J., Westby, G.W.M., Redgrave, P., Rector, D.M., and Harper, R.M. 1996. Cerebral vasomotion: 0.1 Hz oscillation in reflectance imaging of neural activity. *Neuroimage* 4: 183-193.

Mitra, P.P. and Bokil, H.S. 2008. *Observed Brain Dynamics*. Oxford University Press, New York.

Mitra, P.P. and Pesaran, B. 1998. Analysis of dynamic brain imaging data. *Biophysical Journal* 76: 691-708.

Mitra, P.P., Pesaran, B., and Kleinfeld, D. 1999. Analysis of dynamic optical imaging data. in *Imaging: A Laboratory Manual* (ed. R. Yuste, F. Lanni, and A. Konnerth), pp. 9.1-9.9. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Nishimura, B., Schaffer, C.B., Friedman, B., Lyden, P.D., and Kleinfeld, D. 2007. Penetrating arterioles are a bottleneck in the perfusion of neocortex. *Proceedings of the National Academy of Sciences USA* 104: 365-370.

Osol, G. and Halpern, W. 1988. Spontaneous vasomotion in pressurized cerebral arteries from genetically hypertensive rats. *American Journal of Physiology* 254: H28-H33.

Papoulis, A. 1962. *The Fourier integral and its applications*. McGraw-Hill Book Company, New York.

Percival, D.B. and Walden, A.T. 1993. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge University Press, Cambridge.

Pesaran, B., Sornborger, A., Nishimura, N., Kleinfeld, D., and Mitra, P.P. 2005. Spectral analysis for dynamical imaging data. in *Imaging in Neuroscience and Development: A Laboratory Manual* (ed. R. Yuste and A. Konnerth), pp. 439-444. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Prechtl, J.C. and Bullock, T.H. 1994. Event-related potentials to omitted visual stimuli in a reptile. *Electroencephalography and clinical neurophysiology* 91: 54-66.

Prechtl, J.C. and Bullock, T.H. 1995. Structure and propagation of cortical oscillations linked to visual behaviors in the turtle. In *Proceedings of the 2nd Joint Symposium on Neural Computation* (ed. T.J. Sejnowski), pp. 105-114. Institute for Neural Computation, University of California, San Diego and California Institute of Technology.

Prechtl, J.C., Bullock, T.H., and Kleinfeld, D. 2000. Direct evidence for local oscillatory current sources and intracortical phase gradients in turtle visual cortex. *Proceedings of the National Academy of Sciences USA* 97: 877-882.

Prechtl, J.C., Cohen, L.B., Mitra, P.P., Pesaran, B., and Kleinfeld, D. 1997. Visual stimuli induce waves of electrical activity in turtle cortex. *Proceedings of the National Academy of Sciences USA* 94: 7621-7626.

Sengupta, A.M. and Mitra, P.P. 1999. Distributions of singular values for some random matricies. *Physical Review E* 60: 3389-3392.

Shih, A.Y., Friedman, B., Drew, P.J., Tsai, P.S., Lyden, P.D., and Kleinfeld, D. 2009. Active dilation of penetrating arterioles restores red blood cell flux to penumbral neocortex after focal stroke. *Journal of Cerebral Blood Flow and Metabolism* 29: 738-751.

Sornborger, A., Yokoo, T., Delorme, A., Sailstad, C., and Sirovich, L. 2005. Extraction of the average and differential dynamical response in stimulus-locked experimental data. *Journal of Neuroscience Methods* 141: 223-229.

Taylor, A.L., Cottrell, G.W., Kleinfeld, D., and Kristan, W.B. 2003. Imaging reveals synaptic targets of a swim-terminating neuron in the leech CNS. *Journal of Neuroscience* 23: 11402–11410.

Thomson, D.J. 1982. Spectral estimation and harmonic analysis. *Proceedings of the IEEE* 70: 1055-1096.

Thomson, D.J. and Chave, A.D. 1991. Jackknifed error estimates for spectra, coherences, and transfer functions. in *Advances in Spectrum Analysis and Array Processing* (ed. S. Shykin), pp. 58-113. Prentice Hall.

Table 1

Quantity		Units	
Name	**Description**	**Sampled data**	**Computational**
Record length	longest time	$\mathbf{T}$	$\mathbf{N}$
Sample time	shortest time	$\Delta t = \frac{T}{N}$	$\mathbf{1}$
Resolution half-bandwidth	lowest frequency	$\Delta f = \frac{p}{T}$	$W = \frac{p}{N}$
Nyquist frequency ($f_{Nyquist}$)	highest frequency	$\frac{1}{2\Delta t} = \frac{N}{2T}$	$\frac{1}{2}$
Temporal range		$[\Delta t, T]$	$[1, N]$
Spectral range		$\left[-\frac{N}{2T}, \frac{N}{2T}\right]$	$\left[-\frac{1}{2}, \frac{1}{2}\right]$
Time-bandwidth product ($\mathbf{p}$)	$p \geq 1$	$T \cdot \Delta f$	$N \cdot W$

Figure Legends

Figure 1. Basics of Fourier transforms and tapers. (A) Example of the process of tapering data. Top panel shows a time series of sine wave with center frequency $\mathbf{f_0} = \mathbf{T}/10$. Middle panel shows a half-sine taper, defined as $sin\{\pi t/T\}$, and a single Slepian taper with $\mathbf{p} = 1$ and $\mathbf{K} = 1$; the norm of both functions are set to 1. Bottom panel shows the products of the tapers and the time series. **(B)** Magnitude of the Fourier transforms of the untapered data, the data tapered with a single taper, and the data tapered with a single Slepian taper. Also shown is the "ideal" representation with power only in the interval [-1/**T**, 1/**T**] surrounding the center frequency. For the case of no taper, the transform is $\tilde{V}(f) \sim sin\{\pi(f - f_0)/T\}/[\pi(f - f_0)/T]$ and the first zero is at $\mathbf{f} = \pm 1/\mathbf{T}$ relative to $\mathbf{f_0}$, while for the case of the half-sine taper, $\tilde{V}(f) \sim \cos\{\pi(f - f_0)/T\}/\{1 - [2(f - f_0)/T]^2\}$ and the first zero is at $\mathbf{f} = \pm 3/(2\mathbf{T})$ relative to $\mathbf{f_0}$. There is no analytical expression for the transform of the Slepian taper. **(C)** The family of four Slepian tapers, $\mathbf{w^{(k)}(t)}$, for the choice $\mathbf{p} = 2.5$.

Figure 2. Analysis of the intrinsic motion of the diameter of a penetrating arteriole from rat (unpublished data of A. Y. Shih). (A) Two-photon line-scan data through the center of a penetrating vessel over parietal cortex was obtained as described (Shih et al. 2009); $\mathbf{f_{Nyquist}} = 500$ Hz. **(B)** Time series of the diameter as a function of time, as derived from the line scan data as described (Devor et al. 2007) ($\mathbf{T} = 540$ s). The insert shows an expanded region to highlight the multiplicity of rhythmic events present in the signal. **(C)** Spectrum (Eqs. 8 and 9 with $\mathbf{p} = 24$) of the time derivative of the diameter, **D'(t)** (Eq. 11) plotted on log-log (top) and linear-linear (bottom) axes. The gray bands encompass the 95 % confidence bands (Eqs. 13 - 16) and appear symmetric on a log scale (top) but asymmetric on the linear scale (bottom). The frequencies are labeled $\mathbf{f_V}$ for vasomotion, $\mathbf{f_B}$ for breathing, and $\mathbf{f_H}$ for heart beat. The insert in the top figure shows the spectrum of the diameter **D(t)**; note the steep, $\sim 1/f^2$ trend that is removed by taking the spectrum of **D'(t)**.

Figure 3. Analysis of voltage-sensitive dye imaging experiments to find followers of Tr2 (Taylor et al. 2003). (A) Cartoon of the leech nerve cord; input to Tr2 forms the drive signal **U(t)**. **(B)** Fluorescence image of ganglion 10 stained with dye. **(C)** Ellipses drawn to encompass individual cells and define regions whose pixel outputs were averaged to form the optical signals $\mathbf{V_i(t)}$. **(D)** Simultaneous electrical recording of Tr2, *i.e.*, **U(t)**, and optical recordings from six of the cells shown in panel C ($\mathbf{T} = 9$ s), *i.e.*, $\mathbf{V_1(t)}$ through $\mathbf{V_6(t)}$, along with $|\tilde{C}_i(f_{Drive})|$ (Eq. 17 with $\mathbf{p} = 6$) **(E)** Polar plot of $\tilde{C}_i(f_{Drive})$ between each

optical recording and the cell Tr2 electrical recording for all 43 cells in panel C. The dashed line indicates the $\alpha = 0.001$ threshold for significance (Eqs. 24 and 25); error bars one standard error (Eqs. 18 – 25).

Figure 4. Denoising of spinning-disk confocal imaging data on Ca^{2+} waves in organotypic culture (unpublished data of J. T. Vogelstein). (A) Selected frames from a 1200 frame sequence of 100 x 100 pixel data. **(B)** The same data set after reconstruction with 25 of the 1200 modes (Eq. 29). Denoising is particularly clear when the data is viewed as video clips. **(C)** Singular value decomposition of the imaging sequence in panel (A). The spectrum for the square of the eigenvalues for the space and time modes. Note the excess variance in the roughly 25 dominant modes (Eqs. 27 and 28). **(D)** The top 15 spatial modes, $\mathbf{F_n(s)}$, plus high-order modes. Light shades correspond to positive values and dark shades negative values. The amplitude of the modes is set by the orthonormal condition $\sum_{t=1}^{N_t} \mathbf{F_m(t)\, F_n(t)} = \delta_{nm}$. **(E)** The top 15 temporal modes of $\mathbf{G_n(t)}$. The amplitude of the modes is set by the orthonormal condition $\sum_{t=1}^{N_t} \mathbf{G_m(t)\, G_n(t)} = \delta_{nm}$.

Figure 5. Analysis of single-trial voltage-sensitive dye imaging data to delineate collective excitation in visual cortex of turtle (Prechtl et al. 1997). (A) Spectrogram of the response averaged over all active pixels in the imager (Eqs. 30 and 31). **(B)** Space-time response during the period while the animal was presented with a looming stimulus. The data were denoised (Eq. 29), low-pass filtered at 60 Hz, and median filtered (400 ms width) to remove a stimulus-induced depolarization. We show every 8-th frame (126 Hz); note the flow of depolarization from left to right. Insert is the spectrum for the interval 4.7 to 5.7 s. Insert is the power spectrum over the $\mathbf{T} = 1$ s interval that encompasses this epoch (black band in panel A). **(C)** Coherence, $\hat{C}(f_0)$, over intervals both before and after the onset of ($\mathbf{T} = 3$ s; $\mathbf{K} = 7$) estimated at successive frequency bins; $\hat{C}(f_0) > 0.14$ indicates significance (Eqs. 33 - 35). **(D - G)** Spatial distribution of amplitude (red for maximum and blue for zero) and phase ($\pi/12$ radians per contour; arrow indicates dominant gradient) of the coherence at $\mathbf{f_0} = 3$, 8, 18, and 22 Hz, respectively, during stimulation.

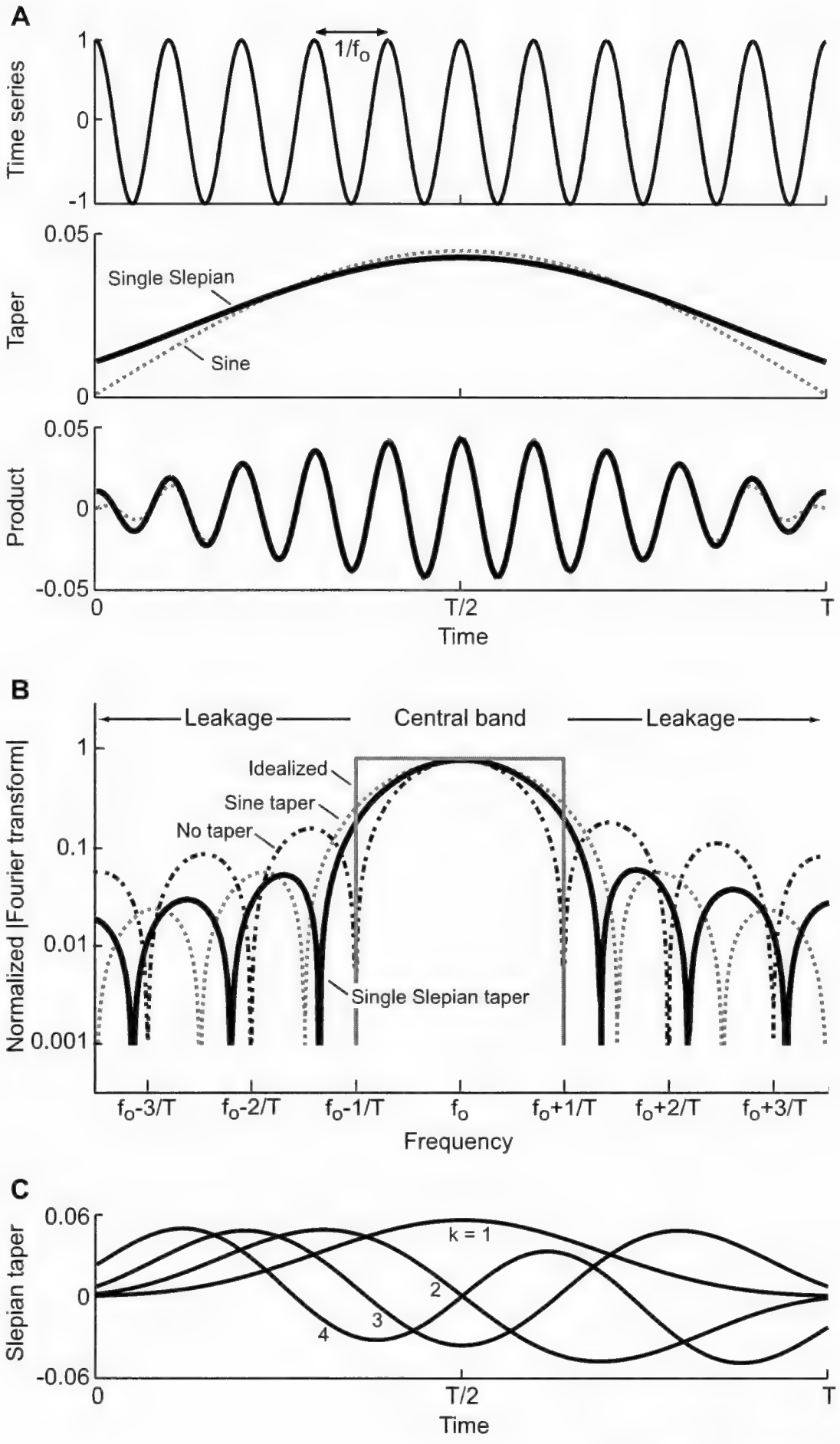


Figure 1 - Kleinfeld and Mitra

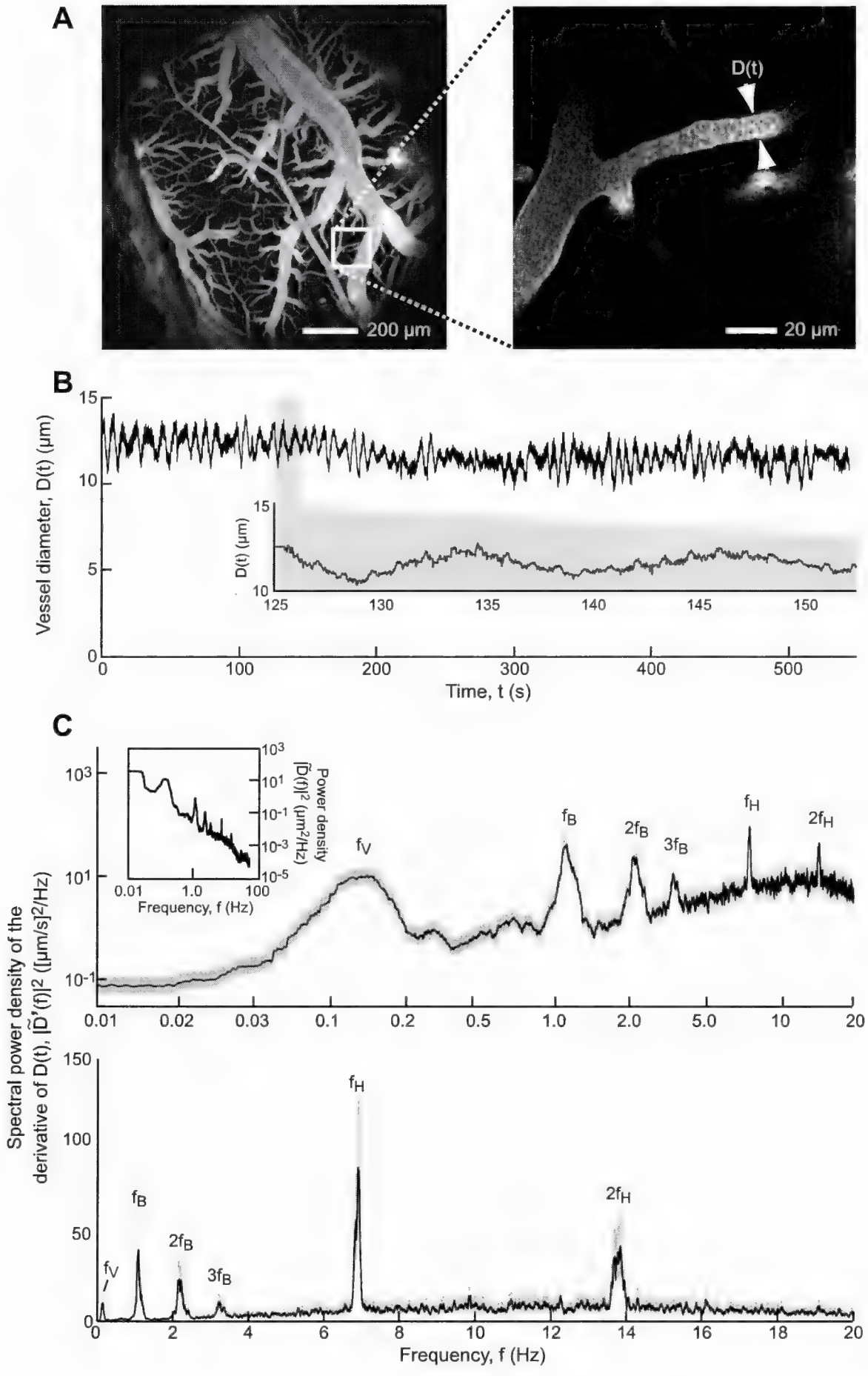


Figure 2 - Kleinfeld and Mitra

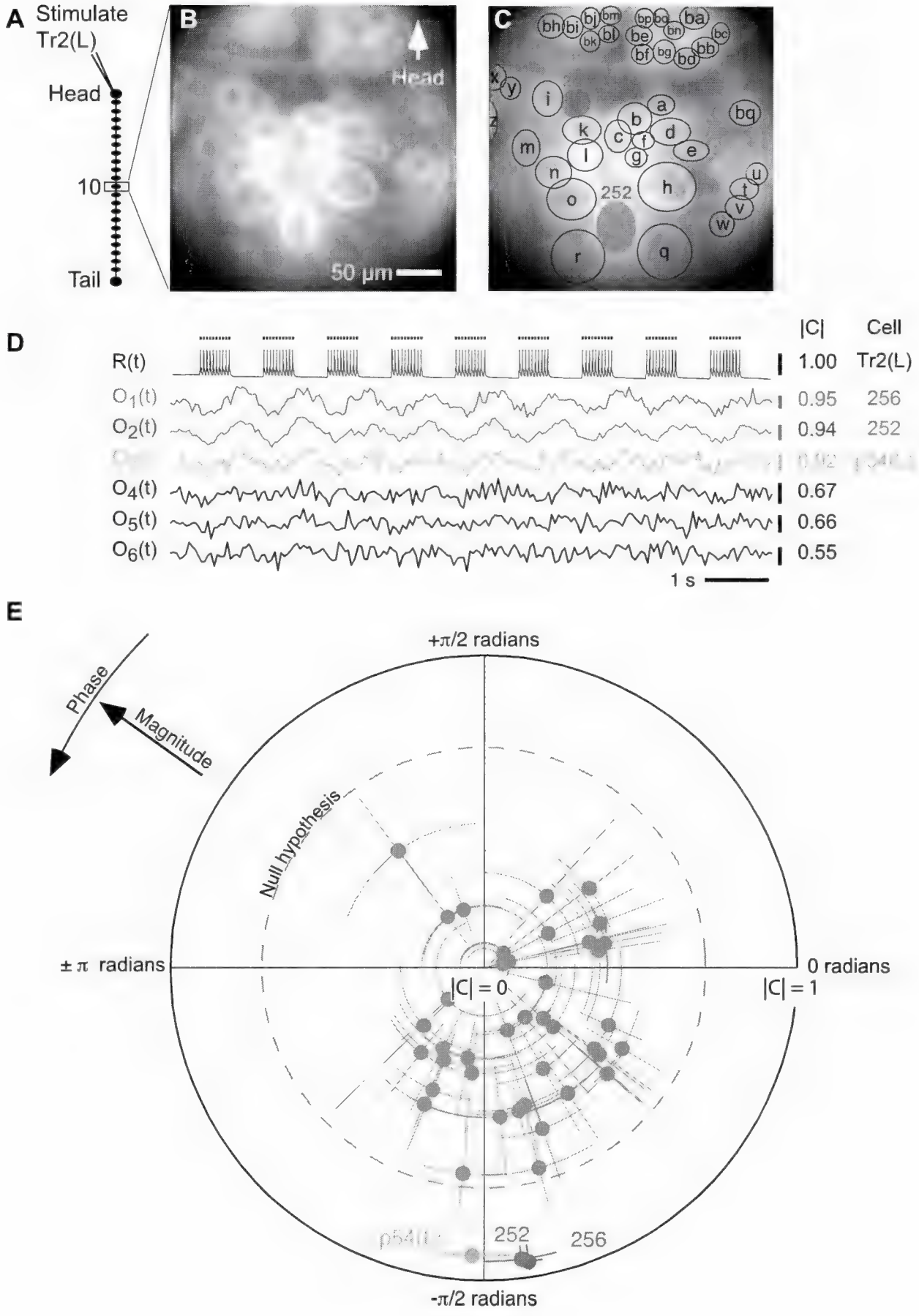


Figure 3 - Kleinfeld and Mitra

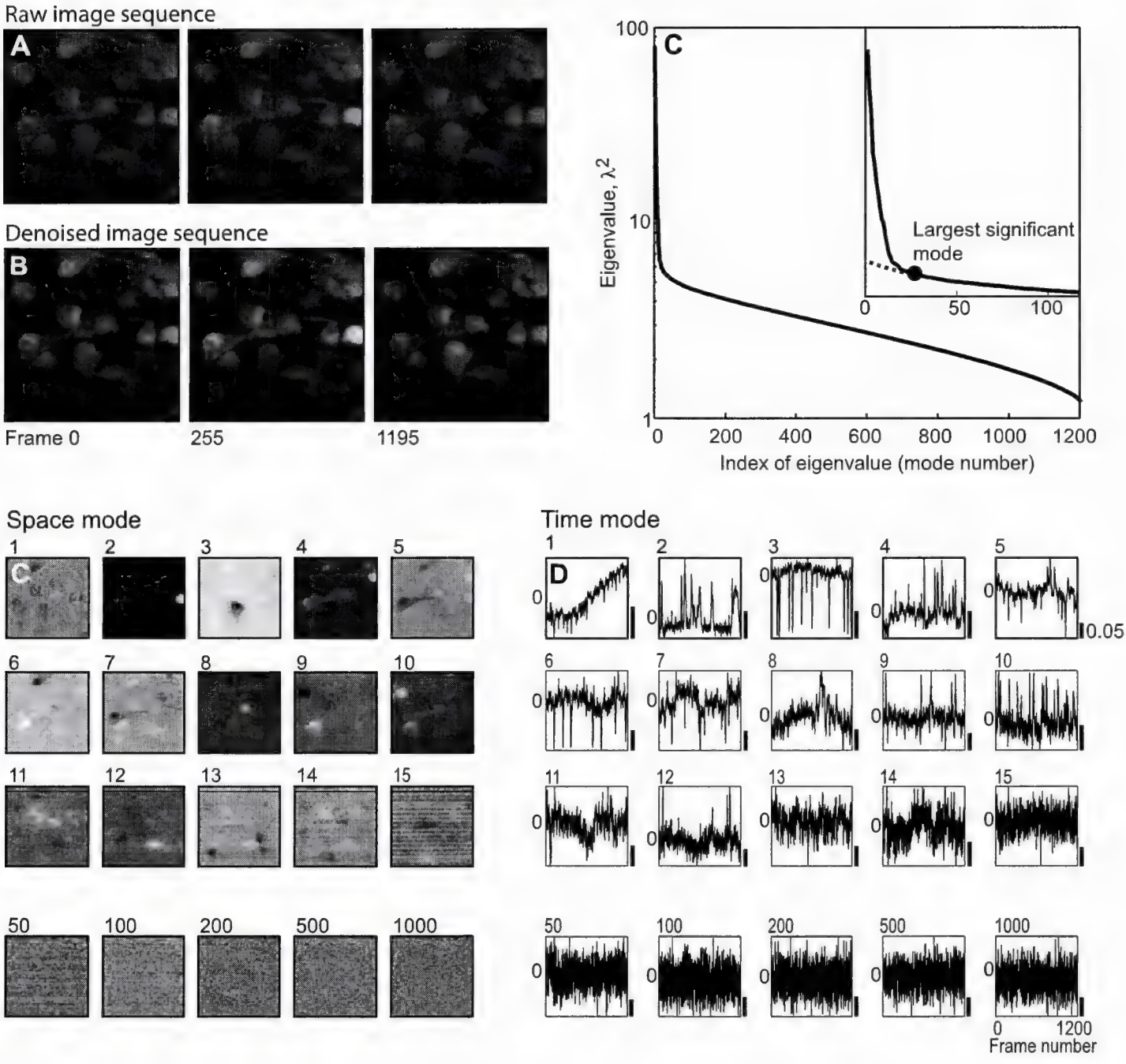


Figure 4 - Kleinfeld and Mitra

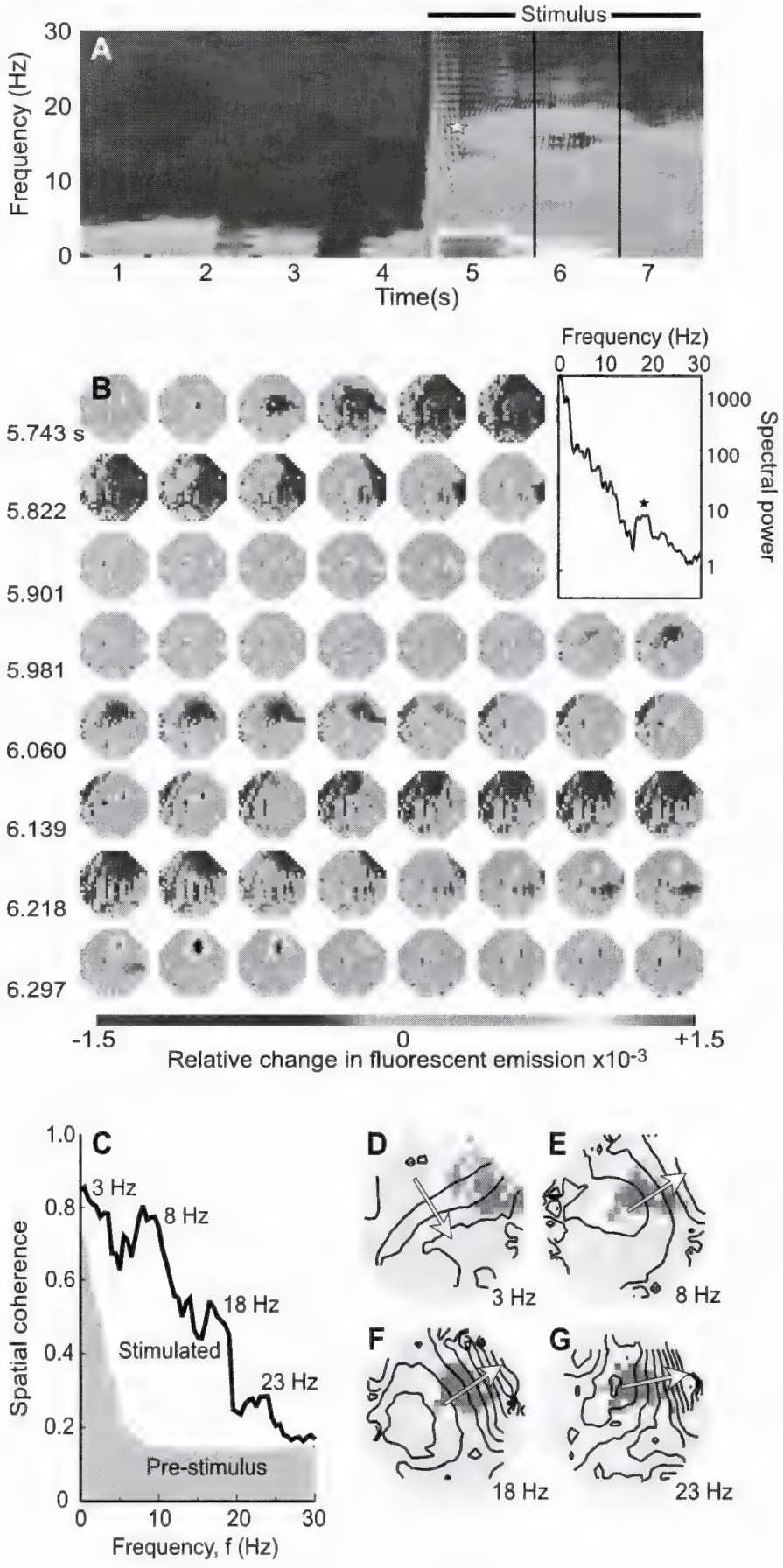


Figure 5 - Kleinfeld and Mitra

Modeling Single Neurons Using the NEURON Simulation Environment

MCN 2009

Tutorial Scope

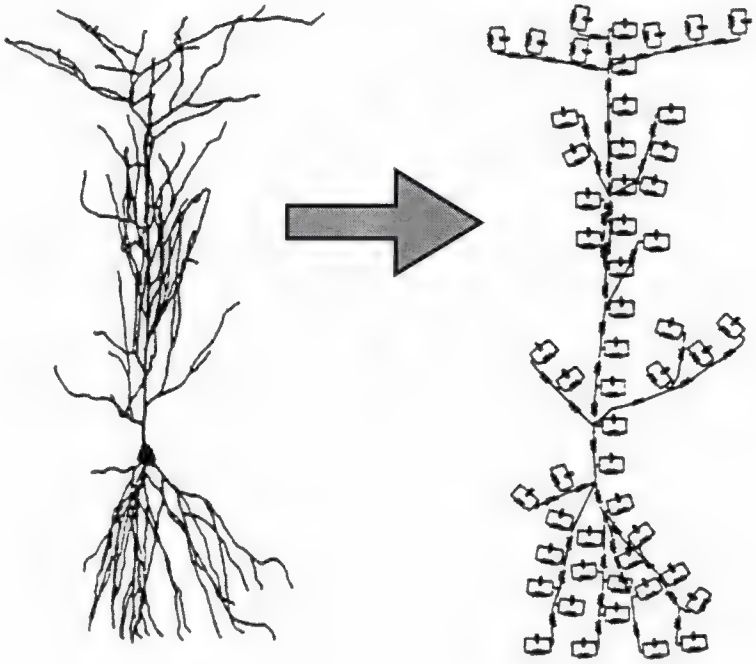
1. A (very) brief overview of the model building process.
2. Do a few simple experiments.
3. Start working on more involved/open-ended questions.

What this tutorial is NOT about

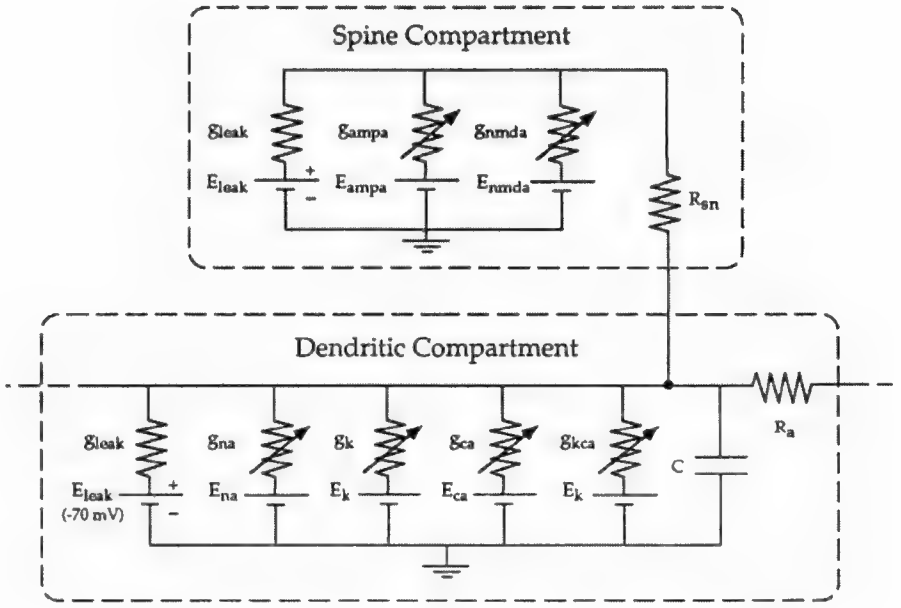
- Making you NEURON pros.
- The NEURON folks will to do that in a one week course (Warning: $$).
- The NEURON Book (TNB) is a good reference. (Cheaper)
- http://www.neuron.yale.edu/neuron/ http://www.neuron.yale.edu/neuron/sfn97/ntc1/outline.htm (Cheapest)

How NEURON views a neuron?

Neuron Electrical Circuit



1 Dendrite +
1 Spine Compartment

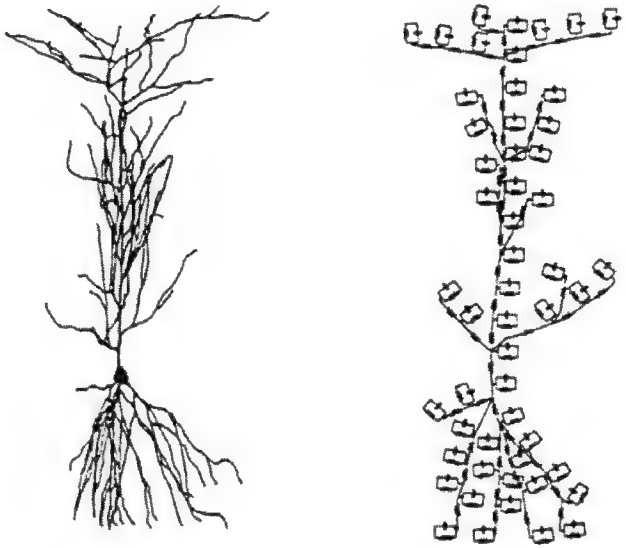


Zador et al. (1992)

•How NEURON views a neuron?

Sections vs. Segments

neurites



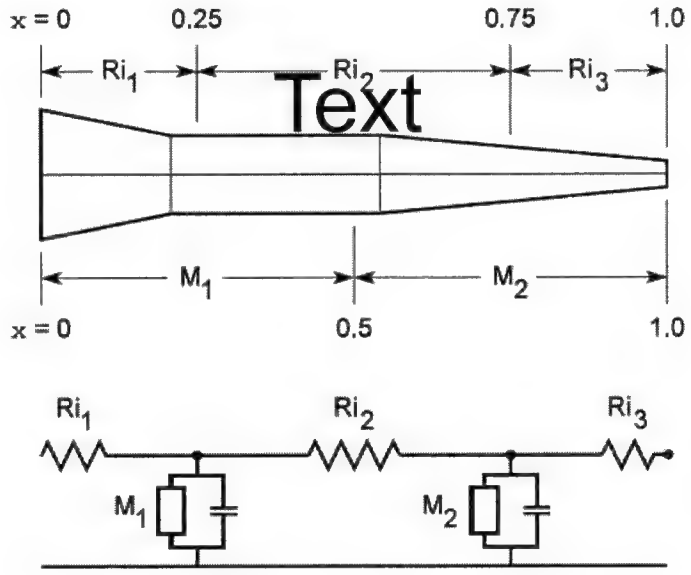
# of nodes

Zador et al. (1992)

•How NEURON views a neuron?

Sections vs. Segments

neurites



Hines & Carnevale, 1997

# of nodes

How many segments to model the dendrite “perfectly”? (TNB, Chapter 5.7)

Building a single neuron model: Step 1: Morphology

➡ Topology: how are the different sections connected?

- (TNB, Chapter 5, Section 5.5) GUI: CellBuilder

➡ Geometry: what is the shape of each section. Couple of ways to do this:

- length & diameter of the section.

- a list of 3D co-ordinates & corresponding diameters for the section. GUI: CellBuilder

- (TNB, Chapter 5, Sections 5.6).

  - Hines & Carnevale. 1997

Building a single neuron model: Step 2: Biophysics

- Distributed Mechanisms: HH Na^+ channel density, passive membrane leak, etc.
  Example units: S/cm^2 GUI: CellBuilder

- Point Processes: Synapses and electrodes.
  Example units: μS, nA GUI: CellBuilder

(TNB, Chapter 5, Section 5.6)

NMODL: A model description language for user-defined biophysical properties (TNB, Chapter 9). GUI: ChannelBuilder

Execution

- Add instrumentation: specify signals to monitor.
- Set-up simulation run controls (simulation time, time step, numerical integration methods).
- Run!

- (TNB, Chapter 7)

What does NEURON provide

- NEURON interpreter "nrniv": hoc command interpreter

- NEURON GUI: specify morphology, biophysical properties, instrumentation, execution control

- nrnivmodl (UNIX), makenrndll (OS X and WINDOWS) - links custom mechanisms into the NEURON executable.

NEURON File Types

- .ses

- Saves the morphology + biophysical properties + instrumentation + run control as setup by the NEURON GUI.

NEURON File Types

- .hoc

- Morphology (`create`, `section diam`, `section L`, `pt3add`, `connect`)
- Biophysical Properties (distributed/ point processes)
- Execution Control (list `hoc` commands, `proc`)
- Instrumentation (`plot`, `Shape`)

NEURON File Types

- .mod
  - Define Biophysical Mechanisms (distributed/point processes)

NeuronDB
http://senselab.med.yale.edu/NeuronDB/
ModelDB
http://senselab.med.yale.edu/modeldb/

NEURON Help

- Quick Reference Manual
- http://www.neuron.yale.edu/neuron/docs/help/quick_reference.html
  - HOC commands
  - Model Description Language (NMODL) syntax

Helpful HOC commands

- `topology()` - information on topology of the currently loaded morphology.
- `psection()` - information for current accessed section.
- `access` - makes section the default currently accessed section.

Exercises

NEURON Exercises for MCN course

Topic: dendritic integration

Overview

These exercises are meant to accompany Mel's morning lecture on Wednesday, August 12.

Exercise 1: Signal propagation in a passive thin dendrite

In this exercise, we will explore how signals propagate asymmetrically in a thin (basal) dendrite towards and away from the cell body. One first goal will be to replicate the results shown in Figure 4 of Nevian, Larkum, Polsky & Schiller (2007). We will focus in this exercise on passive (i.e. voltage-independent) synapses and dendrites, contrasting the effects of fast vs. slow synaptic conductance inputs. We will also observe the dramatic difference between current and voltage attenuation in the same branch.

NMDA channels make things more complicated (and interesting) because they are voltage dependent, and will be considered in Exercise 2.

A NEURON model and GUI will be provided, based on the reconstructed morphology a layer 5 pyramidal neuron. The point process "AmpaNmdaSynapse" will be used to model synaptic events as alpha functions; you will modify only the time course parameter (tauAMPA) and the peak conductance value gmax (gmaxAMPA). Leave the NMDA parameters set to zero, we will use these in Exercise 2.

Procedure

1. Launch NEURON. Double click on file mbl.hoc to launch NEURON.

2. Determine the peak conductance of the synapse. Using the PointProcessGroupManager, place a synapse at the most distal point on the dendrite (the dendrite you will start with is about 200 µm long). Run the simulation a few times, varying the peak conductance until you get a 0.1 mV peak EPSP at the soma. Make a note of the conductance, which you will use at varying locations on the branch.

3. Collect raw voltage (attenuation) data. Place the synapse at 5-10 roughly equally-spaced locations along the dendrite, beginning at the soma (or close to soma) and ending at the most distal point. Watch the voltage and space plots during the simulations to get an intuition for how voltage signals travel inward toward the soma and outward toward the tip. For each stimulus location, record in a spreadsheet (1) the distance from soma, and the peak voltage both at the (2) site of stimulation and (3) at the soma. You can read these values interactively from the voltage plots using the mouse pointer, and you can check them against the values on the "max" space plot. Hint: Read off the distance to the soma from the space plot.

4. Plot the attenuation data. Using a spreadsheet of your choice (e.g. Excel/Calc/GDoc), first make a combined plot as a function of dendritic location (as is done in Figure 4a in the Nevian et al paper) that includes: (1) the peak depolarization at the stimulus site, and (2) the peak depolarization at the soma. In a second plot (or on the same plot), plot of the ratio of the above 2 plots as a function of position.

Questions: *(1) How large an attenuation factor can you get for an inward-travelling signal? (2) How large an attenuation can you get for outbound signals? (3) What fraction of the current injected at a distal location makes it to the cell body? (4) How can you reconcile 1 and 3? (5) What functional interpretation do you give to this strange asymmetric pattern of attenuations? (6) Does your pattern of attenuation look like that in the Nevian et al. paper? Why or why not? FYI, a branch that is a perfect 1 um cylinder has been included with the morphology.*

5. In preparation for looking at the effects of NMDA conductances in the next exercise, repeat the above exercise (steps 2-4) for a much slower synaptic conductance, by increasing tau to, say, 10 ms.

Question: *Before doing the experiment, how do you think the pattern of attenuation will change, if at all? Hint: scale the synaptic conductance to keep the somatic depolarization roughly matched to your earlier plot.*

Exercise 2: Signal propagation when dendritic spikes are involved

Here we will look at a similar issue of signal propagation, but when the excitatory synapses include NMDA channels. To do this, set a non-zero time constant and conductance for the NMDA component of the channel.

Procedure

1. Pick a proximal (near the soma) and a distal (near the tip) location to stimulate the dendrite. Generate separate plots of (1) local and (2) somatic voltage responses to a series of increasingly powerful synaptic inputs at the two sites (4 plots total). Increase the number of synapses activated at each site until you observe an "NMDA spike", that is, where there is a sudden jump in the response, and then go a few steps beyond. To keep the traces from run to run, select "Keep Lines" from the "View" menu. Note the number of synapses you stimulated just below and just above the NMDA spike threshold at the two sites. You will use these 4 numbers below.

Question: *Are the basic inward and outward attenuation patterns consistent with what you found in Exercise 1 when everything was passive?*

2. Manually read off the peak somatic voltages from your plots for the proximal and distal stimulation cases, and plot these values as a function of the number of activated synapses. Both cases can go on the same plot in Excel in different colors.

Question: *What is the same about the two plots, and what is different? Can you predict this from what you know of the input resistance? Is the dendrite "democratic", meaning that all synapses have the same effect on the soma?*

3. Now pull 8 values off of these plots and make a column graph to compare peak somatic responses generated by proximal vs. distal stimulation at 4 stimulus levels (i.e. keep the same stimulus level at the

two sites in each of these 4 cases): (1) just below distal NMDA spike threshold; (2) just above distal NMDA spike threshold; (3) just below proximal NMDA spike threshold; and (4) just above proximal NMDA spike threshold. For each case, you should have 2 columns, one for the distal stimulus site, and one for the proximal.

Question: *Looking at your bar graphs, can you make general statements about the relative potency of proximal vs. distal stimulation, as measured by the size of the somatic response the inputs cause? Can you make functional sense of this pattern of results?*

Exercise 3: Implications of varying Rm

In the previous exercises, there was an implicit assumption that dendrites have uniform membrane resistance along their length. (For simplicity, we are lumping together different channels that contribute to the resting membrane properties, including Ih). This assumption has been challenged in some experimental studies (Magee 1998, Magee and Cook 2000). An interesting case to consider is one where Rm falls off moving away from the cell body to keep the input resistance roughly constant as a function of position in the branch. Would this somehow make the branch act more like an idealized uniform cable with symmetric current and voltage attenuation in both directions? Use the GUI to manipulate Rm, and see if you can determine what the functional effects would be by repeating some of the above experiments.

Instructions for manipulating Rm via NEURON GUI.

1. In the menu choose Build->Cell builder.

2. Go to "Management" tab, select "Import".

3. Click "Import", select "Top Level"; If warning box appears, select "Go ahead and import".

4. Check 'Continuous Create' to synchronize changes between the cell builder tool and the workspace.

4. Go to "Subsets" tab, choose a select method. either one should work and have its own use as the name suggests.

5. Select some section(s) and click "New SectionList", give a name (for example "mysubtreelist" if you use the subtree select mode, "dend16list" if you choose branch dend[16] using the select one mode, or "basals" if you use the basname selection mode) for the SectionList your just created.

6. Click "Parameterized Domain Page", and click "Create a SubsetDomainIterator". You have many choices as to how to specify distance along the dendrite, try the default which is "Path length from the root".

7. Go to "Biophysics" tab and select the SectionList we created in step 5. Make sure "Specify Strategy" is checked, click on the mechanism that you want to manipulate which is "pas" in this case. Now select the iterator you created in step 6 and select the variables in "pas" that need to be manipulated, in this case, "g_pas".

8. Uncheck the "Specify Strategy" box, set e_pas back to -65mV. Now select the iterator g_pas that can be specified inhomogeneously.

9. select a function of f(p), p is the "path distance from root", you may see different functions if you made different choices in step 6. You have three already available functions, and you can type in your own expressions as for loops if "New" is chosen. Specify parameters for the function you chose. Bingo!

Further Investigations

Possible projects include:

1. Major et al. (2008) argued that a thin dendrite functions like a chain of separate subunits. This raises the question: how much independence of voltage signals do you expect to see in a dendrite that is being driven by synapses at different locations? You could investigate this with experiments as in Exercise 2, with two closely spaced inputs rather than a very proximal and a very distal input. Space plots showing the voltage response along the length of the branch through time will be a particularly good way to look at this.

References

Nevian, T., Larkum, M. E, Polsky A., and Schiller J. 2007. Properties of basal dendrites of layer 5 pyramidal neurons: a direct patch-clamp recording study. *Nat Neurosci* 10, no. 2: 206-14.

Stuart, G., and Spruston, N.. 1998. Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. *J. Neurosci.* 18, no. 10 (May 15): 3501-3510.

Magee, JC. 1999. Dendritic Ih normalizes temporal summation in hippocampal CA1 neurons. *Nat Neurosci* 2, no. 6 (June): 508-514.

Magee, JC, 1998. Dendritic hyperpolarization-activated currents modify the integrative properties of hippocampal CA1 pyramidal neurons. *J. Neurosci.*, 18, 7613-7624.

Magee, JC and Cook EP. 2000. Somatic EPSP amplitude is independent of synapse location in hippocampal pyramidal neurons. *Nat Neurosci* 3, no. 9: 895-903.

Major G, Polsky A., Denk W., Schiller J. and Tank DW . 2008. Spatiotemporally graded NMDA spike/plateau potentials in basal dendrites of neocortical pyramidal neurons. *Journal of Neurophysiology* 99, no. 5 (May): 2584-2601.

Short problems for NEURON

Bartlett Mel
MCN 2009

Choose at least one:

1. Katz et al. (Neuron 63, 171–177, July 30, 2009) showed that synaptic weights on apical oblique dendrites of CA1 pyramidal neurons are progressively smaller and sparser moving from the trunk end of the branch towards the distal end. They argued that this pattern is more consistent with a 2-layer model of the neuron, as opposed to the opposite gradient of synaptic weights that they might have found, which they would interpret as supporting a "global" model of integration. Can you do experiments in a model that support or contradict their conclusions?

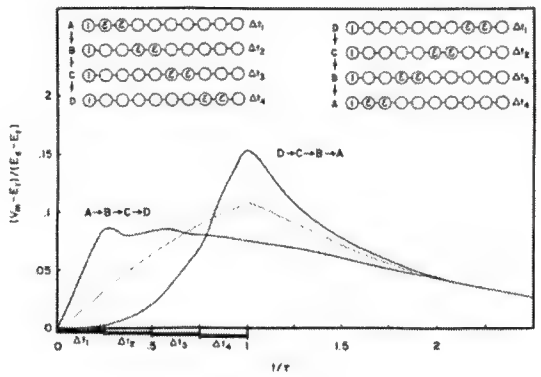
2. Simulate the new variant of the "Rall" model of direction selectivity (DS) as discussed by Michael Hausser. In his talk he focused on the subthreshold (for NMDA spikes) range of branch activation. (1) Try to replicate his DS result. (2) As Mike suggested, try the experiment after blocking NMDA channels and see how large a difference you can get between inward vs. outward sweeps with only AMPA conductances. Can you beat 10%? (3) Try to find a stimulus level and/or number of synapses where the outward sweep generates an NMDA spike but the inward sweep doesn't. If so, does this reverse the direction preference?

3. Kampa and Stuart (J. Neurosci, 26:7424-32, 2006) provide evidence that A-type potassium currents are elevated in distal basal dendrites. Passive cable properties dictate that distal dendrites have significantly elevated input resistance, such that a synaptic input of fixed size would lead to a much larger local depolarization distally compared to proximally. This passive property could mean distal synapses have an "unfair" advantage in triggering local spikes. Test the hypothesis that the elevated A-current counteracts this increased tendency of distal synapses to generate local spikes.

MCN 2009 Student Assignment

Reproduce directional selectivity in a passive dendritic cable according to Wilfrid Rall's original model (Rall W. (1964) Theoretical significance of dendritic trees for neuronal input-output relations. In Neural Theory and Modeling, ed. R.F. Reiss. Stanford Univ. Press).

The original figure (Figure 7):



Student Assignment:

1. Reproduce the results of the original Rall model (see below for guidelines)

2. Make the conditions more realistic:
   - add a soma
   - use biexponential synaptic conductances
   - shorten the cable: at what length does the effect drop below 5%?

3. Reproduce in a pyramidal cell model with NMDA conductances

Guidelines for reproducing Rall's original simulation of DS:

1. Cable with 10 segments of 1 μm diameter and 0.2 λ each
2. 8 synapses, one on each compartment, from compartments 2-9
3. Synapses are square conductance pulses of 0.25 τ duration
4. syanpses are activated sequentially in pairs at 0.25 τ intervals
5. Max synaptic conductance equals $1/R_m$
6. $R_m = 5000$ and $R_i = 50$
7. Record voltage at compartment 1

Neuron
Report

Synapse Distribution Suggests a Two-Stage Model of Dendritic Integration in CA1 Pyramidal Neurons

Yael Katz,[1,2] Vilas Menon,[3] Daniel A. Nicholson,[1,5] Yuri Geinisman,[5] William L. Kath,[1,3,4] and Nelson Spruston[1,*]
[1]Department of Neurobiology and Physiology
[2]Interdepartmental Biological Sciences Program
[3]Department of Engineering Sciences and Applied Mathematics
[4]Northwestern Institute on Complex Systems
Northwestern University, Evanston, IL 60208, USA
[5]Department of Cell and Molecular Biology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA
*Correspondence: spruston@northwestern.edu
DOI 10.1016/j.neuron.2009.06.023

SUMMARY

Competing models have been proposed to explain how neurons integrate the thousands of inputs distributed throughout their dendritic trees. In a simple global integration model, inputs from all locations sum in the axon. In a two-stage integration model, inputs contribute directly to dendritic spikes, and outputs from multiple branches sum in the axon. These two models yield opposite predictions of how synapses at different dendritic locations should be scaled if they are to contribute equally to neuronal output. We used serial-section electron microscopy to reconstruct individual apical oblique dendritic branches of CA1 pyramidal neurons and observe a synapse distribution consistent with the two-stage integration model. Computational modeling suggests that the observed synapse distribution enhances the contribution of each dendritic branch to neuronal output.

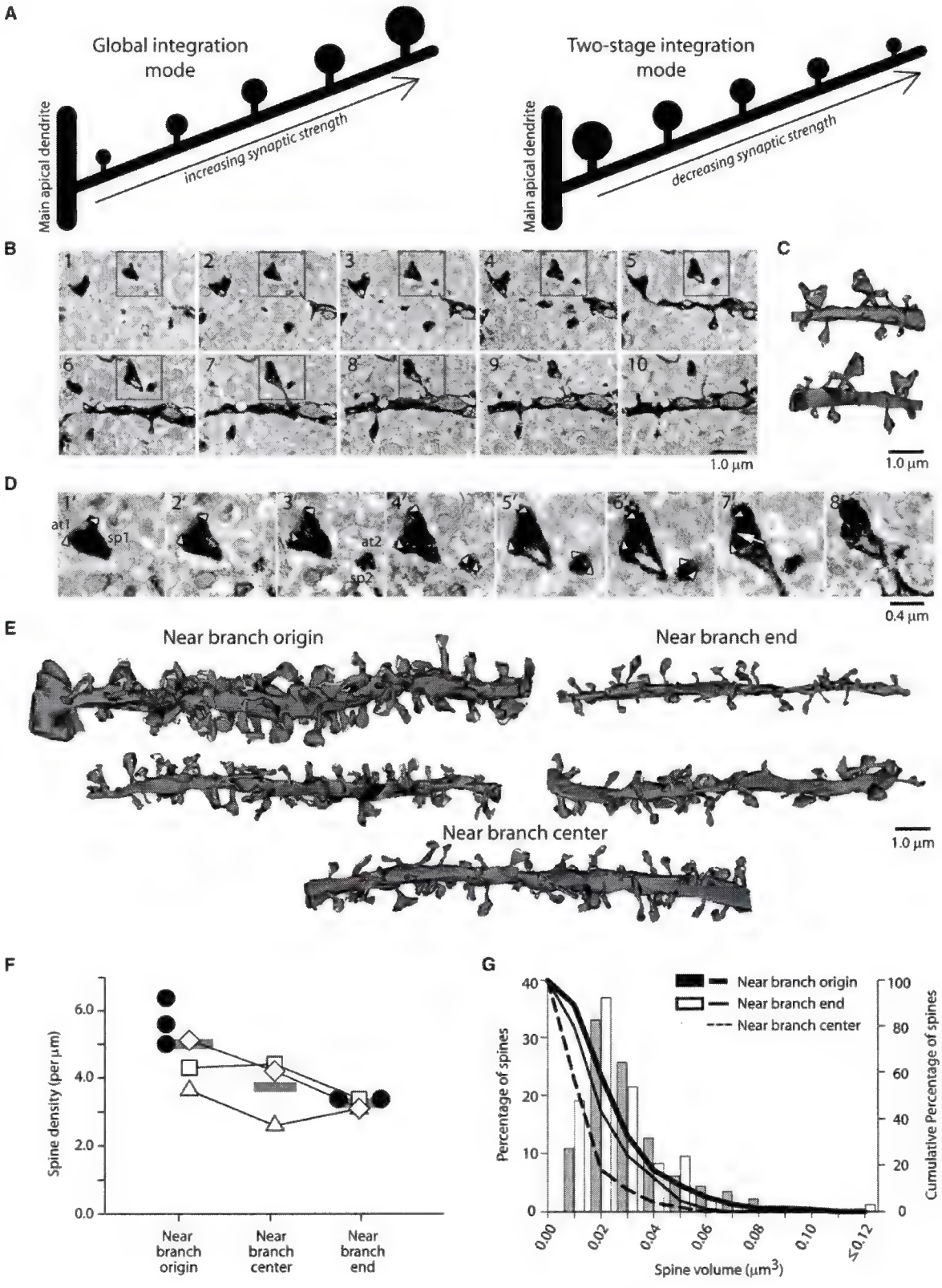
INTRODUCTION

Pyramidal neurons have extended apical and basal dendritic trees, covered with synapses that collect inputs from their network partners. Understanding how these inputs are integrated into all-or-none action-potential output is fundamental to understanding how circuits function in the cerebral cortex. As other work progresses toward understanding the electrophysiological and behavioral implications of different synaptic inputs (Brun et al., 2002, 2008) as well as the molecular basis of the function and plasticity of individual synapses (Harvey and Svoboda, 2007; Harvey et al., 2008; Tanaka et al., 2008), it is equally important to understand how these inputs are integrated at the cellular level and how this integration is shaped by plasticity (Losonczy et al., 2008). Because pyramidal neurons often span millimeter distances and have tens of millimeters of branching dendritic cable (Amaral and Witter, 1989), a crucial question is whether all inputs contribute equally to neuronal output or are weighted by their locations (Häusser, 2001).

Pyramidal neurons can operate in two integration modes (Gasparini and Magee, 2006). In the global integration mode, inputs contribute directly to neuronal output by triggering excitatory postsynaptic potentials (EPSPs) that propagate to the action potential initiation zone in the axon. The axon integrates the contribution from each input, and if the resulting voltage is above threshold, an action potential is generated. In the two-stage integration mode, inputs activate voltage-gated channels in the dendrites and trigger dendritic spikes (Gasparini et al., 2004; Golding and Spruston, 1998). Output from each branch, rather than each synapse, then propagates forward to the axon where a global summation takes place (Gasparini and Magee, 2006; Polsky et al., 2004).

Cable theory predicts that the contributions of inputs to the somatic and dendritic membrane potentials will depend on their locations (Rall, 1959; Spruston, 2008). Inputs that arrive close to the soma will have a larger influence on axonal action potential output than more distal inputs, which attenuate more by the time they reach the action potential initiation zone (Golding et al., 2005; Rall, 1959). Distal inputs, however, lead to larger local dendritic voltage changes compared to their more proximal counterparts, because of differences in input resistance along individual dendritic branches. Small-diameter dendritic segments near branch ends tend to have the highest input resistance, which decreases toward the branch origin where the conductance due to larger dendritic branches and the soma is greatest. Thus, proximal inputs will have relatively large effects in the soma but smaller effects in the dendrite, and distal inputs will have relatively small effects in the soma but larger local effects in the dendrite.

Despite the fact that this prediction follows from basic principles of cable theory that have been well understood for at least 50 years, how pyramidal neurons balance these competing effects has never been directly tested. Experiments in CA1 pyramidal neurons show that for limited distances along the somatodendritic axis, neurons normalize the representation of inputs at the soma by increasing synaptic strength to balance attenuation of synaptic current (Magee and Cook, 2000; Nicholson et al., 2006). Whether synaptic conductance is scaled in this manner along individual dendritic branches, however, is not known.


A
Global integration mode
Main apical dendrite
increasing synaptic strength
Two-stage integration mode
Main apical dendrite
decreasing synaptic strength
B
1.0 µm
C
1.0 µm
D
at1
sp1
at2
sp2
0.4 µm
E
Near branch origin
Near branch end
Near branch center
1.0 µm
F
Spine density (per µm)
6.0
4.0
2.0
0.0
Near branch origin
Near branch center
Near branch end
G
Near branch origin
Near branch end
Near branch center
Percentage of spines
Cumulative Percentage of spines
Spine volume (µm3)
0.00
0.02
0.04
0.06
0.08
0.10
≤0.12

This is a critical issue because if inputs are to contribute approximately equally to neuronal output, then the global and two-stage integration modes suggest opposite distributions of synaptic strength along individual dendritic branches. If neurons are optimized for the global integration mode, synapses farthest from the soma should have the largest conductance in order to compensate for charge attenuation, so synaptic conductance should increase from the origin to the end of individual dendritic branches (Figure 1A, left). If neurons instead are optimized for the two-stage integration mode, synapses farthest from the soma should be the weakest, in order to balance the high input impedance near branch ends, and thus synaptic conductance should decrease from the origin to the end of individual dendritic branches (Figure 1A, right).

To ascertain what synapse distribution is actually present in CA1 pyramidal neurons, we used serial-section electron microscopy (ssEM) to reconstruct excitatory synapses along individually labeled apical oblique dendrites. This approach differs from previous ssEM experiments (Megias et al., 2001; Nicholson and Geinisman, 2009; Nicholson et al., 2006), which involved sectioning blocks of tissue to obtain measurements that are an average over many neurons and dendritic locations. By determining the synapse distribution along individual dendrites, we could infer whether the synapse distribution in CA1 pyramidal neurons favors a global or two-stage integration mode.

RESULTS AND DISCUSSION

We injected biotinylated dextran amine (BDA) into the rat hippocampus in vivo to achieve sparse labeling of CA1 neurons (see Experimental Procedures). We then used light microscopy to select isolated neurons with clearly labeled spiny dendrites. By using ssEM (Figures 1B–1D; Figure S1 available online), we reconstructed three apical oblique dendrites in their entirety, and for each branch we made within-dendrite comparisons of the synapses near the branch origin, near the center of the branch, and near the branch end (Figure 1E, bottom). Additionally, we reconstructed three proximal and two distal dendritic segments from different branches, identified by the branch point with the primary apical dendrite and the termination of the dendritic branch, respectively (Figure 1E, top), and we performed between-dendrite comparisons of these segments. Analysis was restricted to dendritic segments at least 10 μm long and spines that were completely contained within the serial sections. A total of more than 3000 sections were analyzed. In all cases, spines were completely filled and traceable to their parent dendrite, and in many cases postsynaptic densities (PSDs) were clearly visible (spine, n = 433 proximal and n = 189 distal; PSD, n = 195 proximal and n = 78 distal; Figures 1B–1D; Figure S1). Spine volume and PSD area measurements were within the range reported in other studies (Harris et al., 1992; Harris and Stevens, 1989; Nicholson and Geinisman, 2009; Nicholson et al., 2006).

We found two major differences between spines on dendritic segments near branch origins and those near branch ends. First, spine density was approximately 50% greater in near-branch-origin segments compared to those near branch ends (t test, $p < 0.003$; Figure 1F). Second, the distribution of spine volumes was skewed toward larger spines in dendritic segments near branch origins ($\chi^2 = 20.4$; $p < 0.05$; Figure 1G).

We also compared synapses near branch origins to those near branch ends. Though we found a variety of synapse shapes and sizes at all dendritic locations (Figure 2A), the distribution of PSD areas was heavily skewed toward larger PSDs near branch origins ($\chi^2 = 354.9$; $p < 0.00001$; Figure 2B). This trend was present in the overall analysis (Figure 2B; mean PSD area 0.057 μm^2 near branch origins and 0.030 μm^2 near branch ends), as well as in each within-dendrite comparison (Figure S2).

To control for the possibility that synapses near the terminal portion of the dendrite are unusual, we also analyzed 101 spines and synapses in three dendritic segments from the middle of the branch. We found that in each of these segments, spine density and the distributions of spine volume and PSD area were more similar to the distributions near the dendrite's terminal end than near its origin (Figures 1E–1G and 2B; Figure S2), suggesting that the decline in these values is not restricted to the most terminal portion of dendritic branches. Spine volume near the center of the branch appeared to be smaller than near branch ends; the significance of this result will require further study.

To quantify the relationship between spine volume or PSD area and synaptic strength (Kharazia et al., 1996; Kharazia and Weinberg, 1999), we performed immunogold labeling of AMPA receptors (Figures 2C and 2D). We analyzed synapses and their parent spines (n = 342) from area CA1 stratum radiatum of the hippocampus, each of which could be traced back to their spine necks. We found that both spine volume and PSD area were strongly correlated with the number of gold particles (Figure 2E, $R^2 = 0.7216$, $p < 0.001$; Figure 2F, $R^2 = 0.7567$, $p < 0.001$), indicating that these measures provide reasonable estimates of relative synaptic strength. Thus, assuming that PSDs on different dendritic segments have similar receptor properties,

Figure 1. Contrasting Models of Synaptic Integration/Spines near and far from the Primary Apical Dendrite

(A) Schematic showing the predicted distribution of synaptic conductance along an individual apical oblique dendritic branch if synapses are scaled to normalize somatic EPSP amplitude (left) or the probability of initiating a dendritic spike (right).

(B) Serial-section electron microscopic view of a segment of dendrite with spines and synapses.

(C) Three-dimensional reconstruction of the segment shown in (B). Dendrite is shown in gray, spines are shown in purple, and postsynaptic densities (PSDs) are shown in blue. Top panel is the reconstruction in its orientation as shown in the electron micrographs. Bottom panel is the reconstruction rotated 180°.

(D) Higher-magnification serial electron micrographs of the boxed regions in (B). Electron micrographs show two dendritic spines (sp1 and sp2) making synapses with presynaptic axon terminals (at1 and at2). The borders of the PSDs are marked by white arrowheads and the full arrow indicates a perforation.

(E) Three-dimensional reconstructions of segments of the same dendrite (bottom three segments) and different dendrites (top two segments).

(F) Scatter plot showing spine densities in dendritic segments near branch origins, near branch ends, and near branch centers. Open symbols represent segments from the same branch, filled symbols represent segments from different branches, and lines show the means.

(G) Histogram showing the relative and cumulative frequencies of spine volumes in dendritic segments near branch origins, near branch ends, and near branch centers.

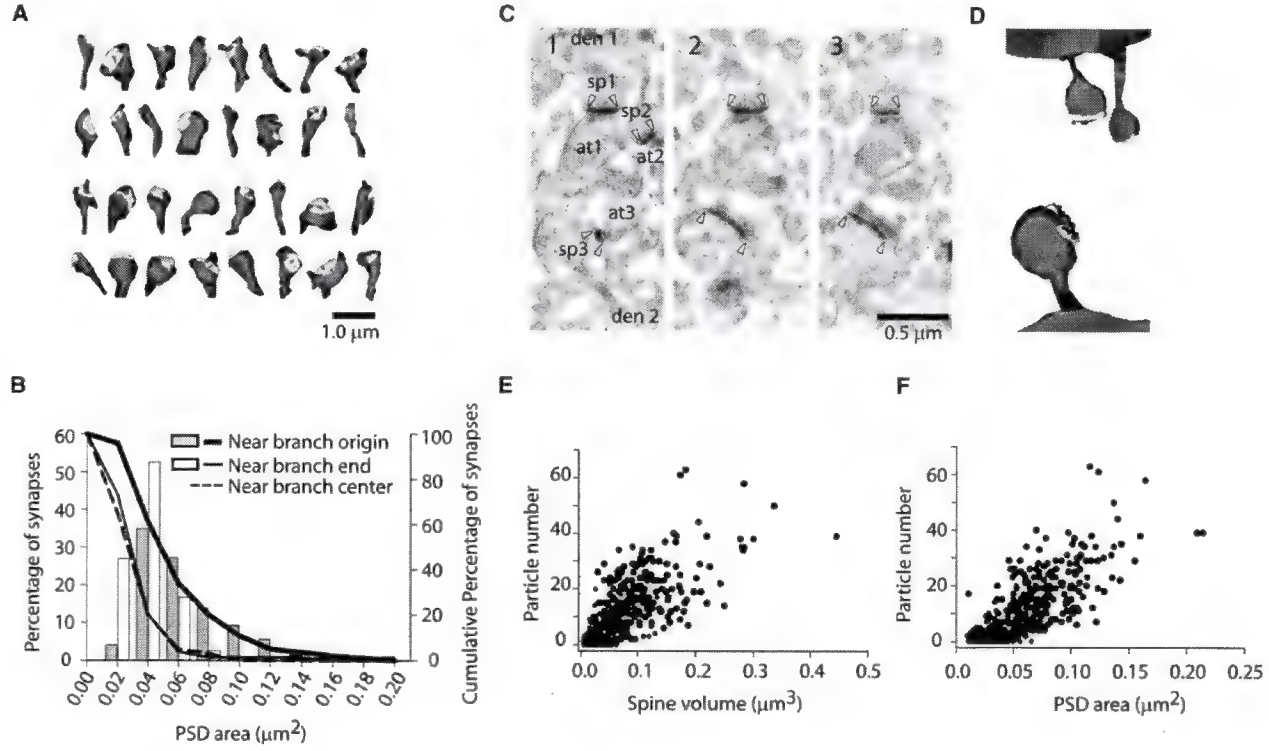


Figure 2. Synapses near to and far from the Primary Apical Dendrite

(A) Reconstructions of several spines and synapses.

(B) Histogram showing the relative and cumulative frequencies of PSD areas in dendritic segments near branch origins, near branch ends, and near branch centers.

(C) Electron micrographs of serial sections showing postembedding immunogold labeling of AMPA receptors.

(D) Three-dimensional reconstruction of synapses (blue), their spines (purple), and the synaptic immunogold particles (black) shown in (C).

(E) Correlation of particle number with spine volume.

(F) Correlation of particle number with PSD area.

the approximately 2-fold reduction in PSD size at the branch end relative to the branch origin suggests that near-branch-end synapses are about half as strong as their near-branch-origin counterparts. From these data, we conclude that synapse strength is scaled within each dendrite to reduce the impact of location-dependent differences in dendritic depolarization due to varying input impedance along the branch.

To investigate the functional implications of the synapse distribution that we observed, we incorporated our experimental findings into a computational model. By using a reconstructed CA1 pyramidal neuron morphology with excitable dendrites (see Experimental Procedures; Figure 3A), we distributed synapses along each apical oblique branch according to the experimentally determined gradient, with synapse density and strength both decreasing from branch origin to branch end in a step-like manner after the first third of the branch. Groups of synapses throughout each branch were randomly selected (from the actual distribution of PSD areas; see Experimental Procedures) and activated, often leading to the initiation of dendritic spikes (60–80 mV amplitude; Figure 3B). Spikes tended to initiate near the end of each branch, attenuating severely by the time they reached the primary apical dendrite and even further between this location and the soma, consistent with experimental data (Losonczy and Magee, 2006). The number of activated synapses required to trigger a dendritic spike varied because synapse locations and hence strengths were chosen randomly, so simulations were repeated to determine average responses to activation of different populations of synapses on the same branch (500 trials for each branch). The average number of activated synapses needed for a dendritic spike varied between branches as well, but across all branches an average of approximately 10 synapses triggered a dendritic spike, again consistent with experimental data (Losonczy and Magee, 2006).

We performed a comparison between this experimentally based model and the hypothetical model that would be predicted if synapses were instead distributed to normalize the contribution of inputs to global integration in the axon. In this hypothetical "global integration model," synapse density was uniform and synaptic conductance was chosen to approximately normalize the somatic depolarization produced by inputs along each apical branch (keeping the total synaptic conductance the same as in the experimentally based model). We then repeated the same simulations. If synaptic conductance is indeed scaled to normalize the contribution of inputs to dendritic spike initiation, we would expect

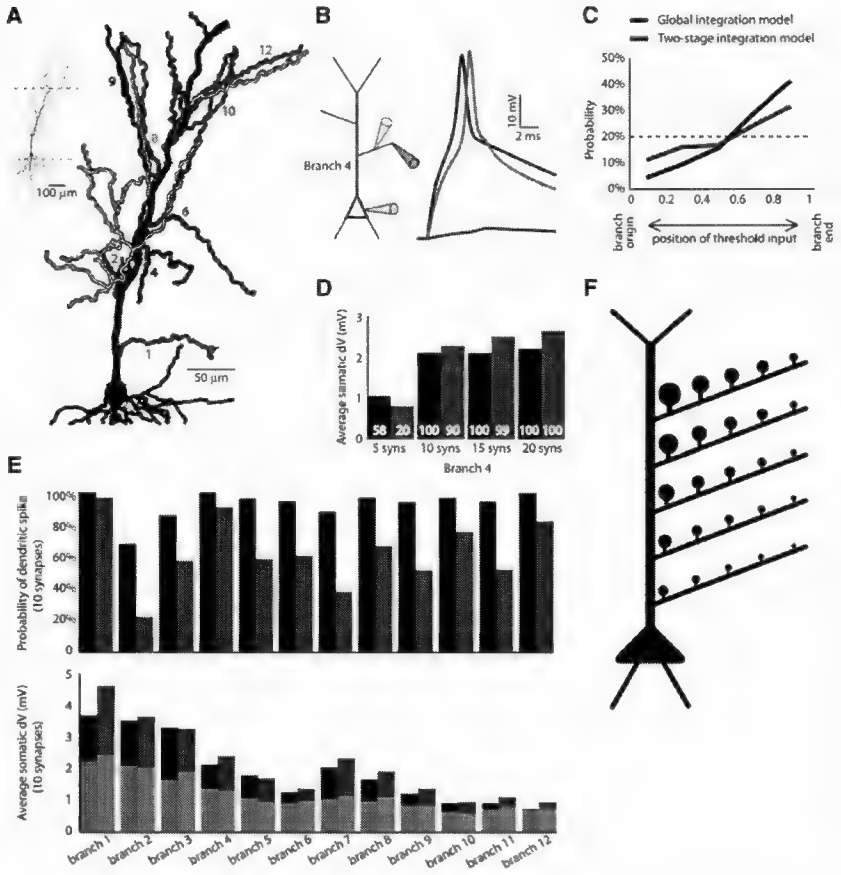


Figure 3. Functional Consequences of Synapse Distribution

(A) The reconstructed CA1 pyramidal neuron morphology used for the models and the apical oblique dendrites used for the simulations (region between dashed lines enlarged).

(B) Sample voltage traces from a simulation in which 15 synapses were randomly selected and activated on branch 4 in the two-stage integration model (188 µm from the soma). Voltage is indicated at the branch end (black trace), center of the branch (red trace), branch origin (green trace), and soma (blue trace). Synapses were distributed along each branch shown in (A) with decreasing density and strength from branch origin to branch end as observed experimentally (two-stage integration model) and the reverse (global integration model).

(C) For each branch, synapses were randomly activated until a dendritic spike occurred (200 trials/branch). Probability distribution (bin size = 0.2) showing the location along the dendritic branch of the input triggering the spike (the last synapse selected) for the two models (black and blue lines). The dashed line indicates the uniform distribution expected if the final input contributes equally to the spike at all locations.

(D) Somatic depolarization averaged over 500 trials resulting from activation of 5, 10, 15, and 20 synapses in the two models: global integration (black bars) and two-stage integration (blue bars). The white numbers indicate the percentage of trials that resulted in a dendritic spike.

(E) Top: The percentage of trials resulting in a dendritic spike for the global (black bars) and two-stage (blue bars) integration models when ten synapses are activated on each branch (500 trials per branch). Bottom: The average somatic depolarization resulting from these simulations shown separately for trials in which a dendritic spike was triggered (dark bars) and trials that did not produce a dendritic spike (light bars).

(F) Schematic of the proposed synapse distribution for the CA1 apical dendritic tree based upon previous (Nicholson et al., 2006; Magee and Cook, 2000) and current results.

the location of the last input selected—the one that brings the branch above spike threshold—to be uniformly distributed (Figure 3C, dashed line). In the global integration model, inputs near branch ends were approximately 9.4 times more likely to trigger a dendritic spike as inputs near branch origins (Figure 3C, black). In our experimentally based model, inputs near branch ends were only 2.7 times as likely to bring the branch above dendritic-spike threshold as those near branch origins (Figure 3C, blue). Because of the high input resistance associated with branch ends, this means that dendritic spikes are often triggered by small numbers of inputs near branch ends in the global integration model, a situation greatly reduced in the model based on the actual synapse distribution. Thus, the experimentally observed distribution helps to normalize the likelihood for synapses at different positions to contribute to a dendritic spike.

A second important consequence of the synapse distribution we observed is that it increases the contribution of a dendritic branch to axonal action potentials (Figure 3D; Movies S1–S6). Despite the fact that dendritic spikes occurred more frequently in the model predicted by global synaptic integration, when fixed numbers of synapses on a branch were activated in the two models, the depolarization measured at the soma tended to be greater in the model based on the actual synapse distribution (Figures 3D and 3E). The global integration model produced a larger response only when a small number of synapses were activated, leading to a much higher likelihood of dendritic spike initiation owing to the prevalence of strong synapses on the high-impedance terminal portions of the branch (Figure 3D). The enhanced representation of each branch at the soma in the two-stage integration model was attributable primarily to the larger numbers of stronger synapses close to the branch origin. By contrast, in the global integration model (or even with a uniform distribution of synaptic density and strength; not shown), synaptic weight is shifted more distally and thus the output of each branch is reduced.

Our model predicts that the ability of individual dendritic branches to influence axonal output should decrease with the distance of the branch origin from the soma and axon, because both EPSPs and dendritic spikes attenuate through the leaky dendritic cable (Figure 3E). Our model, however, did not account

for the fact that distal branches on average contain more powerful synapses, which may boost the representation of these branches at the soma (Magee and Cook, 2000; Nicholson and Geinisman, 2009; Nicholson et al., 2006) and thus minimize the distance dependence of branch output. Combining the results of the present study with previous results (Magee and Cook, 2000; Nicholson and Geinisman, 2009; Nicholson et al., 2006), we propose that synapse strength increases along the somato-dendritic axis but decreases along the length of each radial oblique branch (Figure 3F). Additionally, the contribution of the dendritic spike from some branches could be enhanced by ion channel gradients that shift the initiation zone to more proximal portions of the branch or by broadening the dendritic spike.

Although asynchronous input may lead to global integration (Gasparini and Magee, 2006), our data suggest that synapses located more distally on a branch will contribute less to neuronal output during this integration mode. During more synchronous synaptic input, however, dendritic branches are likely to generate spikes (Gasparini and Magee, 2006). Our ssEM and modeling results show that synapses are scaled in the direction of normalizing the contribution of individual inputs to dendritic spikes, increasing the output of a branch when a group of synapses trigger a local dendritic spike and thus pointing to the importance of these spikes as fundamental units of synaptic integration in CA1 pyramidal neurons. The capacity to perform multilayer computations enhances the computational power of individual neurons in a network (Polsky et al., 2004; Spruston and Kath, 2004), so it will be important to determine whether other types of neurons employ a similar integration scheme.

EXPERIMENTAL PROCEDURES

Institutional Approval

All experiments were performed with protocols approved by the Northwestern University Institutional Animal Care and Use Committee.

Electron Microscopy

Adult rats (6 months old) were anesthetized with an injection of ketamine and xylazine. Biotinylated dextran amine (BDA-3000, Invitrogen; 10% dissolved in 0.12 M phosphate-buffered saline; PBS) was injected into the subiculum (from bregma: AP = −2.5, ML = +1.5; DV = −4.1) with either pressure injection (0.05 μl/min for 5 min) or iontophoresis (5 μA, alternating at 10 s on/off for 6 min). After 5 days of recovery from surgery, rats were perfused with a fixative containing 3.5% paraformaldehyde, 0.1% glutaraldehyde, and 15% (v/v) picric acid in PBS. Brains were removed, hemisectioned, postfixed in fixative, rinsed in PBS, and then cut into 60 μm-thick slices using a vibratome.

Individual slices were then rinsed in PBS, cryoprotected in ascending concentrations of sucrose and glycerol in PBS, freeze-thawed over liquid nitrogen three times, treated with 1% sodium borohydride in PBS, incubated in 0.3% hydrogen peroxide, and then rinsed. Slices were then rinsed in blocking solution (0.5% BSA, 1% nonfat dry milk, and 5% normal goat serum) and incubated in avidin-biotinylated horseradish peroxidase (HRP) with the Vectastain ABC Elite Kit (Vector Laboratories, PK-6100 Standard, 1:300) for 36 hr at 4°C. Slices were then rinsed thoroughly and the BDA-HRP complex was visualized with diamino benzidine as chromogen (Sigma, SK-4100) under gentle agitation for 10–120 min. Slices containing isolated individually labeled CA1 pyramidal neurons were then rinsed, treated with 0.67% osmium tetroxide and 1% uranyl acetate, dehydrated in graded ethanols, infiltrated with Araldite (Araldite 502), and cured in a drying oven at 60°C for 48 hr for conventional serial section electron microscopy.

Slices were trimmed to isolate the labeled dendrites of CA1 neurons, cut into 68-nm-thin serial sections (100–500 serial sections) with an ultramicrotome (UCT, Leica), and mounted onto gold-gilded nickel slotted grids. Grids were counterstained with 3% aqueous uranyl acetate and Reynold's lead citrate and then mounted in an electron microscope (JEOL 100CX) to photograph serial sections containing dendritic segments within CA1 stratum radiatum in their entirety. Dendritic segments near branch ends were considered to be those segments that could be followed through >50 serial sections and then disappear in subsequent serial sections. Branch-origin dendritic segments were readily identifiable due to their bifurcation from the primary apical dendrite. Second-order dendritic segments (i.e., those connecting a branch off the primary apical dendrite with a daughter terminal dendritic segment) were excluded from analyses.

Electron micrograph negatives were scanned with a PowerLook 2100XL scanner, organized into image stacks for each dendritic segment, and then analyzed and reconstructed with ImageJ (http://rsb.info.nih.gov/ij) and Reconstruct (Fiala, 2005), respectively. Postsynaptic density (PSD) area and spine volume were measured with the same methods as previously described (Nusser et al., 1998; Racca et al., 2000; Nicholson et al., 2006; Nicholson and Geinisman, 2009). In brief, for each spine and each synapse, all PSD or spine profiles on the electron micrographs of serial sections were measured in ImageJ. PSD area was estimated as the product of the summed PSD length and the section thickness (estimated to be 68 nm). When present, perforations or cytoplasmic areas of the spine head not containing a PSD profile were not included in the measurements. Spine volume was estimated as the product of the summed spine surface area values and section thickness.

Postembedding Immunogold Electron Microscopy

AMPA-type receptor immunoreactivity was assessed as previously described (Nicholson et al., 2006). Synapses in CA1 stratum radiatum were included in the analysis only if their host spine could be followed to its spine neck (n = 342). Because of the low-electron density of freeze-substituted tissue, it was impossible, except in fortuitous cases, to follow spines to their parent dendrite. Importantly, however, our measurements are within those reported previously for synapses in CA1 stratum radiatum (Harris et al., 1992; Nusser et al., 1998).

Computational Modeling

The CA1 pyramidal neuron morphology used in all simulations was reconstructed from a rat hippocampal pyramidal neuron described previously (Golding et al., 2001). All simulations were performed with the NEURON simulation environment (Hines and Carnevale, 1997).

The model included membrane capacitance and resistance as well as the active conductances sodium (Na), delayed rectifier potassium (K_{DR}), and A-type potassium (K_A). Passive properties were constrained by electrophysiological recording of voltage attenuation from the soma to dendrite in the reconstructed neuron (Golding et al., 2001), and the active conductances Na and K_{DR} were uniformly distributed. K_A in the primary apical dendrite was distributed based on experimental data (Hoffman et al., 1997), and K_A in oblique dendrites was uniformly distributed with the value at the point of connection with the primary apical dendrite. The model will be made available on the authors' website (http://www.northwestern.edu/dendrite).

Fast excitatory synaptic conductances were modeled as a difference of exponentials with a rise time constant of 0.2 ms, decay time constant 5.0 ms, and reversal potential of 0 mV. Dendritic spikes were defined as a voltage exceeding a threshold of −35 mV.

For the simulations in Figure 3 involving the two-stage integration model, synapses were randomly drawn from a distribution starting at 5 synapses/μm from the branch origin through the first third of each branch and then dropping to 3.3 synapses/μm until the branch end. Synaptic weights were determined as follows: We randomly drew from the probability distributions in Figure 2B to determine the PSD area as a function of the location of the synapse. For the two-stage integration model, if the selected synapse was on the proximal third of the branch, we drew from the measured distribution for dendritic segments near branch origins; if it was on the distal third of the branch, we drew from the measured distribution for dendritic segments near branch ends; and if it was on the middle third of the branch, we drew from the measured distribution for dendritic segments near branch centers. Once we determined the PSD area, we chose the corresponding particle number from Figure 2F. The synaptic weight was the particle number times

0.025 nS, which led to values for unitary synaptic conductance between 0.07 and 0.99 nS. For the global integration model, synapses on each branch were distributed uniformly at a density chosen to match the number of synapses per branch in the two-stage integration model. Synaptic conductance increased approximately linearly from 0.21 to 0.33 nS along each branch, leading to a similar amount of somatic depolarization for inputs at all branch locations. Total synaptic conductance was kept the same in both models in order to facilitate direct comparison.

SUPPLEMENTAL DATA

Supplemental Data include two figures and six movies and can be found with this article online at http://www.cell.com/neuron/supplemental/S0896-6273(09)00510-8.

ACKNOWLEDGMENTS

Supported by the National Institutes of Health (NS-46064 to W.L.K. and N.S., NS-35180 to N.S., and AG-17139 to Y.G.) and Northwestern University (Presidential Fellowship to Y.K.). The authors thank Krutika Lakhoo, Annie Liu, and Elizabeth Piekarz for technical support with analysis of EM images and David Ferster, Indira Raman, and Aryeh Warmflash for critical readings of the manuscript. We thank Drs. Kristen Harris and John Fiala for developing and sharing the Reconstruct program.

Accepted: June 23, 2009
Published: July 29, 2009

REFERENCES

Amaral, D.G., and Witter, M.P. (1989). The three-dimensional organization of the hippocampal formation: a review of anatomical data. Neuroscience *31*, 571–591.

Brun, V.H., Otnass, M.K., Molden, S., Steffenach, H.A., Witter, M.P., Moser, M.B., and Moser, E.I. (2002). Place cells and place recognition maintained by direct entorhinal-hippocampal circuitry. Science *296*, 2243–2246.

Brun, V.H., Leutgeb, S., Wu, H.Q., Schwarcz, R., Witter, M.P., Moser, E.I., and Moser, M.B. (2008). Impaired spatial representation in CA1 after lesion of direct input from entorhinal cortex. Neuron *57*, 290–302.

Fiala, J.C. (2005). Reconstruct: a free editor for serial section microscopy. J. Microsc. *218*, 52–61.

Gasparini, S., and Magee, J.C. (2006). State-dependent dendritic computation in hippocampal CA1 pyramidal neurons. J. Neurosci. *26*, 2088–2100.

Gasparini, S., Migliore, M., and Magee, J.C. (2004). On the initiation and propagation of dendritic spikes in CA1 pyramidal neurons. J. Neurosci. *24*, 11046–11056.

Golding, N.L., and Spruston, N. (1998). Dendritic sodium spikes are variable triggers of axonal action potentials in hippocampal CA1 pyramidal neurons. Neuron *21*, 1189–1200.

Golding, N.L., Kath, W.L., and Spruston, N. (2001). Dichotomy of action-potential backpropagation in CA1 pyramidal neuron dendrites. J. Neurophysiol. *86*, 2998–3010.

Golding, N.L., Mickus, T.J., Katz, Y., Kath, W.L., and Spruston, N. (2005). Factors mediating powerful voltage attenuation along CA1 pyramidal neuron dendrites. J. Physiol. *568*, 69–82.

Harris, K.M., and Stevens, J.K. (1989). Dendritic spines of CA 1 pyramidal cells in the rat hippocampus: serial electron microscopy with reference to their biophysical characteristics. J. Neurosci. *9*, 2982–2997.

Harris, K.M., Jensen, F.E., and Tsao, B. (1992). Three-dimensional structure of dendritic spines and synapses in rat hippocampus (CA1) at postnatal day 15 and adult ages: implications for the maturation of synaptic physiology and long-term potentiation. J. Neurosci. *12*, 2685–2705.

Harvey, C.D., and Svoboda, K. (2007). Locally dynamic synaptic learning rules in pyramidal neuron dendrites. Nature *450*, 1195–1200.

Harvey, C.D., Yasuda, R., Zhong, H., and Svoboda, K. (2008). The spread of Ras activity triggered by activation of a single dendritic spine. Science *321*, 136–140.

Häusser, M. (2001). Synaptic function: dendritic democracy. Curr. Biol. *11*, R10–R12.

Hines, M.L., and Carnevale, N.T. (1997). The NEURON simulation environment. Neural Comput. *9*, 1179–1209.

Hoffman, D.A., Magee, J.C., Colbert, C.M., and Johnston, D. (1997). K+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. Nature *387*, 869–875.

Kharazia, V.N., and Weinberg, R.J. (1999). Immunogold localization of AMPA and NMDA receptors in somatic sensory cortex of albino rat. J. Comp. Neurol. *412*, 292–302.

Kharazia, V.N., Phend, K.D., Rustioni, A., and Weinberg, R.J. (1996). EM colocalization of AMPA and NMDA receptor subunits at synapses in rat cerebral cortex. Neurosci. Lett. *210*, 37–40.

Losonczy, A., and Magee, J.C. (2006). Integrative properties of radial oblique dendrites in hippocampal CA1 pyramidal neurons. Neuron *50*, 291–307.

Losonczy, A., Makara, J.K., and Magee, J.C. (2008). Compartmentalized dendritic plasticity and input feature storage in neurons. Nature *452*, 436–441.

Magee, J.C., and Cook, E.P. (2000). Somatic EPSP amplitude is independent of synapse location in hippocampal pyramidal neurons. Nat. Neurosci. *3*, 895–903.

Megias, M., Emri, Z., Freund, T.F., and Gulyas, A.I. (2001). Total number and distribution of inhibitory and excitatory synapses on hippocampal CA1 pyramidal cells. Neuroscience *102*, 527–540.

Nicholson, D.A., and Geinisman, Y. (2009). Axospinous synaptic subtype-specific differences in structure, size, ionotropic receptor expression, and connectivity in apical dendritic regions of rat hippocampal CA1 pyramidal neurons. J. Comp. Neurol. *512*, 399–418.

Nicholson, D.A., Trana, R., Katz, Y., Kath, W.L., Spruston, N., and Geinisman, Y. (2006). Distance-dependent differences in synapse number and AMPA receptor expression in hippocampal CA1 pyramidal neurons. Neuron *50*, 431–442.

Nusser, Z., Lujan, R., Laube, G., Roberts, J.D., Molnar, E., and Somogyi, P. (1998). Cell type and pathway dependence of synaptic AMPA receptor number and variability in the hippocampus. Neuron *21*, 545–559.

Polsky, A., Mel, B.W., and Schiller, J. (2004). Computational subunits in thin dendrites of pyramidal cells. Nat. Neurosci. *7*, 621–627.

Racca, C., Stephenson, F.A., Streit, P., Roberts, J.D., and Somogyi, P. (2000). NMDA receptor content of synapses in stratum radiatum of the hippocampal CA1 area. J. Neurosci. *20*, 2512–2522.

Rall, W. (1959). Branching dendritic trees and motoneuron membrane resistivity. Exp. Neurol. *1*, 491–527.

Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. Nat. Rev. Neurosci. *9*, 206–221.

Spruston, N., and Kath, W.L. (2004). Dendritic arithmetic. Nat. Neurosci. *7*, 567–569.

Tanaka, J., Horiike, Y., Matsuzaki, M., Miyazaki, T., Ellis-Davies, G.C., and Kasai, H. (2008). Protein synthesis and neurotrophin-dependent structural plasticity of single dendritic spines. Science *319*, 1683–1687.

Cellular/Molecular

Calcium Spikes in Basal Dendrites of Layer 5 Pyramidal Neurons during Action Potential Bursts

Björn M. Kampa and Greg J. Stuart
Division of Neuroscience, John Curtin School of Medical Research, Australian National University, Canberra, Australian Capital Territory 0200, Australia

Patch-clamp recording from dendrites has lead to a significant increase in our understanding of the mechanisms underlying signal integration and propagation in neurons. The majority of synaptic input to neurons, however, is made onto small-diameter dendrites, currently beyond the scope of patch-clamp recording techniques. Here we use both calcium and voltage imaging to investigate propagation of action potentials (APs) in fine basal dendrites of cortical layer 5 pyramidal neurons. High-frequency (200 Hz) AP bursts caused supralinear increases in dendritic calcium at distal, but not proximal, basal locations. Supralinear increases in dendritic calcium were also observed at distal basal locations during AP trains above a critical frequency (~100 Hz). Using voltage imaging, we show that single APs undergo significant attenuation as they propagate into basal dendrites, whereas AP bursts lead to generation of dendritic calcium spikes. Focal and bath application of 4-AP increased the amplitude of calcium transients evoked by APs at distal, but not proximal, locations, suggesting that A-type potassium channels regulate AP backpropagation into basal dendrites. Finally, we show that pairing EPSPs with AP bursts is an effective means of activating synaptic NMDA receptors in basal dendrites. The experimental observations on the role of A-type potassium channels in regulation of AP backpropagation in basal dendrites, as well as the generation of dendritic calcium spikes during AP bursts, were reproduced in a morphologically realistic neuronal model with uniform distributions of dendritic sodium, calcium, and potassium channels. Together, these findings have important implications for understanding dendritic integration and synaptic plasticity in cortical basal dendrites.

Key words: action potential; dendrite; neocortex; backpropagation; STDP; imaging

Introduction

Information is represented and transferred in the CNS by fast voltage impulses called action potentials (APs). APs are initiated in the axon and propagate back into the dendritic tree of many neuronal types (Stuart et al., 1997). In cortical pyramidal neurons, this AP backpropagation is supported by dendritic voltage-gated sodium channels, which enable somatic APs to actively propagate into the apical dendritic arbor in a decremental manner (Stuart and Sakmann, 1994). Most studies have focused on AP backpropagation into the larger-diameter apical dendrite of pyramidal neurons, because these dendrites are readily accessible to patch-clamp recordings. The majority of synaptic input to pyramidal neurons, however, is made onto the much finer basal dendrites (Larkman, 1991), currently beyond the scope of patch-clamp recording techniques. Knowledge of AP backpropagation into these small dendrites is incomplete and will be particularly important during interaction of backpropagating APs with basal synaptic input. Basal dendrites of layer 5 pyramidal neurons receive the majority of synaptic input from neighboring pyramidal neurons (Markram et al., 1997b); synapses that have been extensively used to study spike-timing-dependent synaptic plasticity (STDP) (Markram et al., 1997a; Sjostrom et al., 2001, 2003). An understanding of AP backpropagation into basal dendrites is, therefore, critical to understanding the cellular mechanisms underlying STDP at these synapses.

Because of their inaccessibility to patch-clamp recording techniques, recent studies on information processing in basal dendrites of neocortical pyramidal neurons have used fluorescence imaging with calcium-sensitive dyes and voltage-sensitive dyes (VSDs) (Schiller et al., 2000; Antic, 2003; Milojkovic et al., 2004; Polsky et al., 2004). These studies indicate that synaptic inputs onto basal dendrites have the capacity to evoke NMDA spikes (Schiller et al., 2000) and plateau potentials (Milojkovic et al., 2004). In addition, they suggest that basal dendrites support robust AP backpropagation (Antic, 2003). Little is known, however, about the regulation of AP backpropagation during high-frequency AP bursts or the potential role of dendritic calcium electrogenesis in basal dendrites during burst firing. This issue is particularly important given the critical role of AP bursts in information processing (Lisman, 1997; Snider et al., 1998) and synaptic plasticity (Paulsen and Sejnowski, 2000).

Here we investigate the ability of single APs and high-frequency AP bursts to invade the basal dendrites of cortical layer 5 pyramidal neurons. In contrast to previous experimental and

Received July 24, 2005; revised May 26, 2006; accepted May 26, 2006.

This work was supported by funding from the Alexander von Humboldt Foundation and the Australian National Health and Medical Research Council. We thank Lucy Palmer for help with the voltage imaging and Fritjof Helmchen for help with the simulations.

Correspondence should be addressed to Dr. Greg J. Stuart, Division of Neuroscience, John Curtin School of Medical Research, Australian National University, Canberra, Australian Capital Territory 0200, Australia. E-mail: greg.stuart@anu.edu.au.

B. M. Kampa's present address: Department of Neurophysiology, Brain Research Institute, University of Zurich, CH-8057 Zurich, Switzerland. E-mail: kampa@hifo.unizh.ch.

DOI:10.1523/JNEUROSCI.3062-05.2006

Copyright © 2006 Society for Neuroscience 0270-6474/06/267424-09$15.00/0

modeling studies (Vetter et al., 2001; Antic, 2003), we find that single backpropagating APs are severely attenuated in the distal basal dendrites of these neurons because of activation of A-type potassium channels. Importantly, APs can still depolarize the distal basal dendrites during high-frequency burst firing after activation of dendritic calcium channels and the generation of dendritic calcium spikes. These findings have important implications for understanding signal processing, as well as the cellular mechanisms underlying STDP, in basal dendrites

Materials and Methods

Wistar rats (3–4 weeks old) were anesthetized by inhalation of isoflurane and decapitated, and parasagittal brain slices (300 μm) of somatosensory cortex were prepared according to guidelines approved by the Animal Ethics Committee of the Australian National University. During recording, slices were perfused with oxygenated extracellular solution [artificial CSF (ACSF)] containing the following (in mM): 125 NaCl, 3 KCl, 1.25 NaH_2PO_4, 25 $NaHCO_3$, 25 glucose, 2 $CaCl_2$, and 1 $MgCl_2$, pH 7.4 (with 5% CO_2 at 35 ± 1°C). Whole-cell current-clamp recordings were made from the soma of layer 5 pyramidal neurons using a current-clamp amplifier [Molecular Devices (Palo Alto, CA) or Dagan (Minneapolis, MN)] with pipettes containing the following (in mM): 135 K-gluconate, 7 NaCl, 10 HEPES, 2 $MgCl_2$, and 2 Na_2ATP, pH 7.2 with KOH. APs were elicited by brief somatic current injection (2 ms, 2–4 nA). Series resistance was corrected for with bridge balance and capacitance neutralization. Data were filtered at 10 kHz and acquired at 20 kHz on a Macintosh computer (Apple Computers, Cupertino, CA). AxoGraph software (Molecular Devices) was used for acquisition and analysis. Statistical significance was tested using the Student's *t* test at a significance level of 5%. Pooled data represents mean ± SEM.

Calcium imaging. Changes in intracellular calcium concentration were imaged using a confocal microscope [LSM 510 (Zeiss, Oberkochen, Germany) or FV300 (Olympus Optical, Tokyo, Japan)] or a CCD camera (NeuroCCD-SM; Redshirt Imaging, Decatur, GA). Neurons were loaded via the recording pipette with the calcium-sensitive fluorescent dye Oregon Green BAPTA-1 (OGB-1) (200 μM; K_d of 170 nM; Invitrogen, Carlsbad, CA) or Oregon Green BAPTA-6F (OGB-6F) (200 μM; K_d of 3 μM; Invitrogen) added to the pipette recording solution, and cells were filled for at least 30 min before recording. Calcium-sensitive dyes were excited using the 488 line of an argon laser (confocal microscope) or via a 475 ± 15 nm bandpass filter (CCD camera), and emitted fluorescence was detected via a dichroic mirror (510 nm) and long-pass filter (505 or 510 nm). Changes in fluorescence during APs and EPSPs were measured with line scans at 0.3–1 kHz (confocal microscope) or from a small region of interest (ROI) taken at 1 kHz frame rate (CCD camera) across basal dendritic shafts with spines. Calcium signals in spines and the adjacent dendrite were pooled. EPSP-evoked calcium transients were evoked by extracellular stimulation using a pipette filled with external solution plus Sulforhodamine 101 (Invitrogen) placed within 10 μm of the imaged dendrite.

Imaging data were analyzed with Igor (WaveMetrics, Lake Oswego, OR) and expressed as the percentage change in fluorescence relative to baseline ($\Delta F/F$) after background subtraction. For calculation of $\Delta F/F$ ratios during three APs relative to one AP, data were only used if the peak of the one AP response was larger than two times the SD of the baseline noise. In experiments using drug applications, data points obtained during control and washout were averaged to correct for changes in baseline fluorescence during drug application. In the presence of 4-AP (4 mM), TTX was applied locally via a patch pipette filled with HEPES-buffered ACSF including both 4-AP (4 mM) and TTX (2 μM). Pressure was applied to the pipette with a syringe, and data were collected during drug application. After release of pressure, TTX was washed out for up to 5 min.

Voltage imaging. Neurons were filled with the voltage-sensitive dyes JPW1114 (Invitrogen) or its variant JPW3028 (synthesized by J. P. Wuskell and L. M. Loew, University of Connecticut, Farmington, CT) as described previously (Antic, 2003), except that a microscope halogen lamp was used as the excitation light source (Olympus Optical). As described by Antic (2003), cells were filled for ~45 min and then repatched ~2 h later with a pipette without dye. During this 2 h period, AP responses returned to close to control conditions (see Fig. 2*A*). Precautions were taken to reduce photodamage. Only a small area (~30 × 30 μm) was illuminated around the ROI (3 × 4.5 μm), which was visualized at high magnification (60× objectives; Olympus Optical) as shown in Figure 2*B*. This increased spatial resolution and the signal-to-noise ratio and ensured that the ROI was in focus. As a consequence, different dendritic sites were imaged separately. Voltage-sensitive dyes were excited using a 520 ± 45 nm bandpass filter, and emitted fluorescence was detected via a dichroic mirror (575 nm) and long-pass filter (610 nm). Data was obtained at 5 kHz frame rate with a CCD camera (Redshirt Imaging) and digitally filtered with a Gaussian filter at 500 Hz after interpolation of data points to achieve a 20 kHz sample rate. To allow comparison of imaged membrane voltage with that recorded intracellularly via the somatic recording pipette, somatic AP voltage recordings were acquired at the same frequency (5 kHz) using the analog input of the camera software and were also filtered at 500 Hz. AP latency was calculated based on the time difference between the peak of the recorded somatic AP and the dendritic AP-evoked fluorescence change divided by the distance from the middle of the soma to the dendritic location. The rise time of AP-evoked fluorescence transients was defined as the time taken to go from 10 to 90% of the peak amplitude. Traces are averages of 10–100 trials, and background illumination was subtracted. Data was stored as text files and analyzed with Igor (WaveMetrics).

Model of basal dendrites. Simulations of AP backpropagation in basal dendrites were performed using NEURON 5.6 (Hines and Carnevale, 1997) using a reconstructed layer 5 pyramidal neuron model adapted from Stuart and Spruston (1998). Intracellular resistivity, R_i, was set to 150 Ω.cm, membrane resistivity, R_m, was set to 15,000 Ω.cm^2, and membrane capacitance, C_m, was set to 1 μF/cm^2. The following active conductances were included: a fast-inactivating sodium and non-inactivating potassium conductance were incorporated at a uniform density of 30 pS/μm^2, except in the axon in which the sodium conductance density was raised to 30,000 pS/μm^2. Throughout the basal dendrites, a fast-inactivating (A-type) potassium conductance was incorporated at a uniform density of 20 pS/μm^2, together with a uniform distribution of a low-threshold (T-type) calcium conductance at a density of 3 pS/μm^2 and a uniform distribution of a high-threshold calcium conductance at a density of 0.7 pS/μm^2. All properties of active conductances were based on those used in a previous model by Schaefer et al. (2003). In addition, we included a calcium buffer and extrusion mechanism (Helmchen et al., 1996) with a buffer binding ratio of 20 and a clearance rate of 1200 s^{-1} as reported for small dendrites (Sabatini et al., 2002). The volume-averaged calcium concentration in each compartment was calculated based on the total calcium influx and the compartment volume. Simulations were performed at a nominal temperature of 35°C.

Results

Backpropagation of single APs versus bursts

The basal dendrites of large layer 5 pyramidal neurons were imaged using a confocal microscope after being filled with fluorescent calcium-sensitive dyes and were identified by their origin at the soma, the presence of spines, and their typical radial branching pattern within layer 5 (Fig. 1*A*). APs were generated by somatic current injection and dendritic calcium influx associated with single APs and multiple AP bursts measured at different basal dendritic sites using the high-affinity dye OGB-1 (Fig. 1*A*). Single APs produced an increase in intracellular calcium that could be detected at all dendritic locations, with similar increases at proximal (18.3 ± 2.5% $\Delta F/F$ peak; <130 μm from the soma; $n = 12$) and distal (17.3 ± 2.7% $\Delta F/F$ peak; >130 μm from the soma; $n = 18$) sites (Fig. 1*B*,*C*). In contrast, high-frequency (200 Hz) AP bursts generated significantly larger responses at distal dendritic locations (Fig. 1*B*,*C*). Similar responses were observed using the low-affinity dye OGB-6F ($n = 10$) (Fig. 1*D*). Comparison of the response to a burst of three APs at 200 Hz relative to

that evoked by a single AP indicated a distance-dependent, highly supralinear increase in calcium influx during AP bursts ($n = 31$) (Fig. 1*E*). These experiments suggest that AP backpropagation into basal dendrites is significantly more effective during high-frequency bursts, as is the case in the distal apical dendrites of layer 5 pyramidal neurons (Larkum et al., 1999; Williams and Stuart, 2000).

Previous work indicates that dendritic calcium spikes are generated in the distal apical dendrites of layer 5 pyramidal neurons when the frequency of AP trains exceeds a critical value, usually ~100 Hz (Larkum et al., 1999). This "critical frequency" is defined as the frequency of somatic AP firing in which the time integral of the recorded dendritic voltage becomes supralinear. An accompanying supralinear increase in intracellular calcium is observed at distal apical locations as a result of activation of dendritic voltage-activated calcium channels (Larkum et al., 1999). Similar findings were observed in distal basal dendrites, with a supralinear increase in dendritic calcium concentration at a similar critical frequency (Fig. 1*F*,*G*). Trains of five APs at various frequencies lead to a similar increase in calcium at proximal dendritic locations (<130 μm from the soma) (Fig. 1*G*). At distal sites in the basal dendrites (>130 μm from the soma), however, dendritic calcium influx was significantly larger during AP trains at frequencies above 100 Hz (Fig. 1*I*–*G*). On average, at distal dendritic locations (>130 μm from the soma), calcium transients were fivefold (±1.2; $n = 4$; $p < 0.01$) larger during AP trains at supra-critical AP frequencies (>100 Hz) compared with the same number of APs evoked at 66 Hz (Fig. 1*G*). By analogy with previous findings in the apical dendrites of layer 5 pyramidal neurons, these results suggest that summation of APs in distal basal dendrites promotes activation of voltage-gated calcium channels, leading to the generation of calcium spikes.

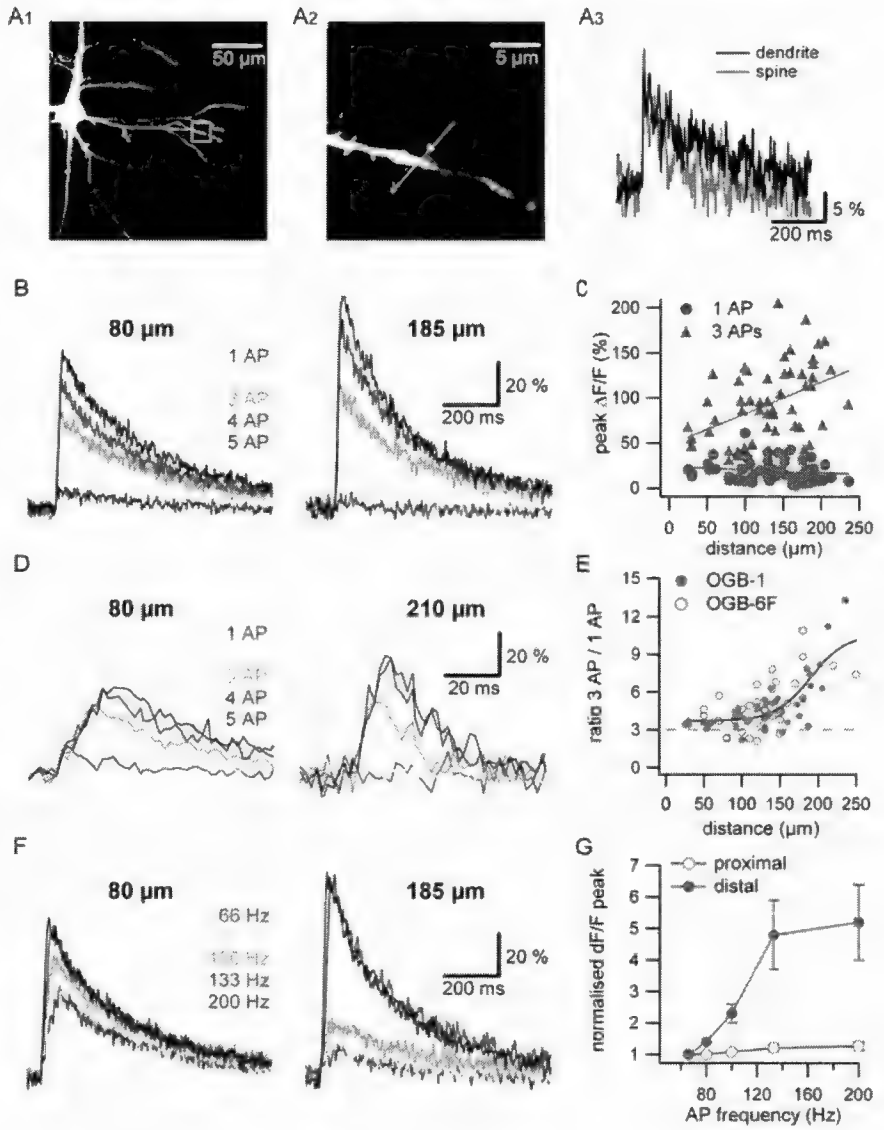


Figure 1. AP bursts evoke supralinear calcium rises in basal dendrites. ***A1***, Confocal image of a layer 5 pyramidal neuron filled with OGB-1. ***A2***, Distal region of basal dendrite indicated by rectangle in ***A1***. Line shows location of line scans. ***A3***, Single APs evoke similar changes in intracellular calcium in the spine (red) and dendritic branch (black). ***B***, Changes in intracellular calcium in response to one to five APs (at 200 Hz) in basal dendrites filled with OGB-1 at the indicated locations. ***C***, Peak $\Delta F/F$ during one (red) or three (green) APs plotted against the distance of the recording site from the soma. Lines show linear fits to the data. ***D***, Calcium imaging of the response to one to five APs (at 200 Hz) in basal dendrites filled with OGB-6F at the indicated locations. ***E***, Ratio of the response to three APs relative to one AP at different distances from the soma in experiments using OGB-1 and OGB-6F. Black line shows sigmoidal fit. Dashed line indicates a linear ratio. ***F***, Calcium transients detected with OGB-1 at proximal and distal sites during trains of five APs at the indicated frequencies. ***G***, Summary data of AP-evoked increases in calcium during trains of five APs at different frequencies. At a critical frequency above 100 Hz, AP trains induced significantly more calcium influx in distal compared with proximal basal dendrites. Data are normalized to changes in calcium at an AP frequency of 66 Hz.

Voltage imaging of dendritic electrogenesis

In the absence of the ability to directly measure membrane potential, we used VSDs to investigate membrane potential changes in basal dendrites during AP bursts (Fig. 2*A*–*C*). A recent study using this method concluded that there is full propagation of single backpropagating APs throughout the entire basal dendritic tree (Antic, 2003). However, because it is not yet possible to quantify the magnitude of changes in membrane potential detected using VSDs, the absolute amplitude of the signals detected is unknown. What is possible is to compare the size of one signal relative to that of another detected at the same location. During backpropagation of AP bursts, we observed little change in the amplitude or shape of the individual APs that makes up the burst at proximal basal sites (<130 μm from the soma) (Fig. 2*D*, gray traces). In contrast, at distal dendritic sites (>130 μm), the amplitude of the third AP in a burst was significantly larger than that of the first (Fig. 2*D*, green traces), with the peak amplitude of the third AP in the burst being on average twofold greater than the first (2.11 ± 0.22; $p < 0.01$) (Fig. 2*D*). Distal locations were imaged before proximal sites in approximately half of these experiments. The amplitude of the third AP relative to the first was not significantly different in these experiments compared with when proximal dendritic locations were imaged before distal sites (distal before proximal, mean of 2.22 ± 0.29, $n = 8$; proximal

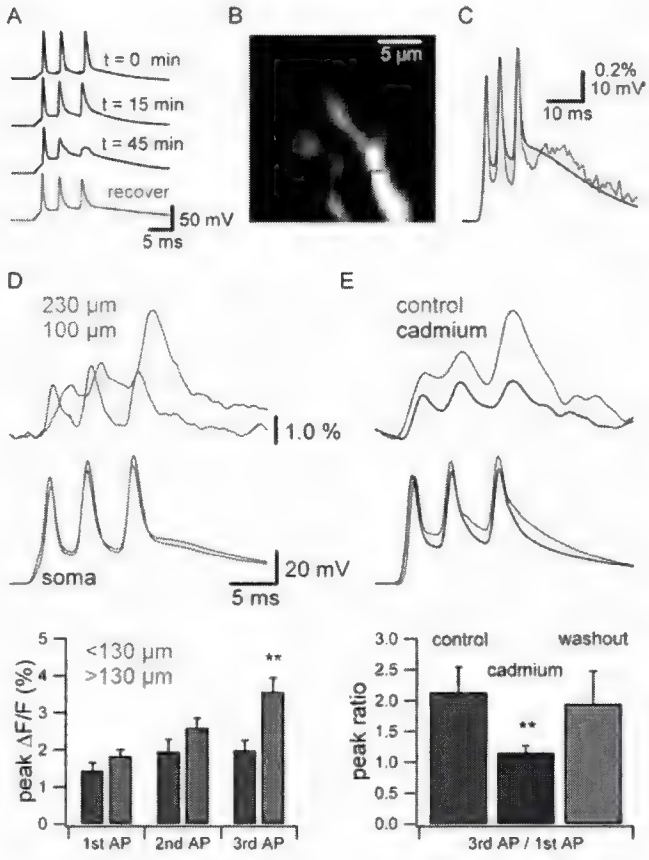


Figure 2. Voltage imaging of AP backpropagation in basal dendrites. ***A***, A three AP burst (200 Hz) recorded at the soma during filling (black traces) and after a 2 h incubation and recovery period (red trace). ***B***, Image of proximal dendrite (average of 30 trials). Rectangle indicates ROI. ***C***, A three AP burst recorded at the soma (black) and imaged in proximal dendrite (red) shown in ***B***. Distance from soma was 30 μm. All data in ***C*** have been filtered at 500 Hz. ***D***, Fluorescence changes in distal (green; 230 μm) and proximal (gray; 100 μm) basal dendrites during a three AP burst (200 Hz). Bottom traces show simultaneously recorded somatic voltage. Traces from different distances were recorded separately. Bottom, Pooled peak $\Delta F/F$ for each AP in a three AP burst at proximal (<130 μm; $n = 6$) and distal (>130 μm; $n = 12$) basal dendritic sites. The peak of the third AP was significantly larger in distal dendrites. $p = 0.005$. ***E***, A three AP burst as in ***D*** in control (gray) and cadmium (200 μM; blue) recorded 200 μm from soma. Bottom, Average ratio of the third to the first AP peak in control ($n = 6$), cadmium ($n = 6$), and after washout ($n = 3$). $p = 0.047$.

before distal, mean of 1.83 ± 0.21, $n = 6$; $p = 0.29$). Given that the region of illumination in these experiments is small (~30 × 30 μm), these data rule out the possibility that photodamage at proximal basal sites is responsible for the observed reduction in amplitude of the third AP relative to the first at distal basal locations.

The most likely explanation of this finding is that, in contrast to the conclusions of Antic (2003), single APs undergo significant attenuation at distal basal dendritic locations, with the third AP in a high-frequency burst boosted as a result of recruitment of dendritic voltage-activated sodium or calcium channels. This explanation would be consistent with the observed supralinear increase in intracellular calcium at distal basal dendritic locations during AP bursts. To test for a role of dendritic calcium channels, we bath applied 200 μM cadmium. Cadmium reduced the peak of the third AP relative to the first (Fig. 2*E*, top traces). On average, at distal basal locations (>130 μm from the soma), the ratio of the third AP relative to the first was close to 1 in the presence of cadmium, and this ratio recovered to control levels after washout (Fig. 2*E*, bottom). These results provide direct evidence that single APs undergo significant attenuation as they backpropagate into distal basal dendrites, whereas AP bursts cause large depolarizations attributable to the generation of dendritic calcium spikes.

Additional evidence for attenuation of AP backpropagation in distal basal dendrites during single APs can be found by an examination of the properties of the first AP in the burst (Fig. 3*A*). The latency between the peak of somatic and dendritic APs was found to increase in a nonlinear manner with distance of the dendritic recording location from the soma (Fig. 3*B*). These data indicate that the conduction velocity of AP backpropagation into basal dendrites is not constant and is slower at more distal dendritic locations. The slowing of AP conduction velocity suggests that backpropagating APs undergo significant filtering as they propagating into basal dendrites. Consistent with this idea, the first AP in a burst showed a more than twofold increase in rise time (10–90%) with recording distance from the soma that saturated at distal dendritic sites (Fig. 3*C*). These results are inconsistent with the idea that single APs propagate in a nondecremental manner into the distal basal dendrites. On the contrary, the observed slowing in AP velocity and increase in AP rise time is consistent with filtering of backpropagating APs by the passive cable properties of the distal basal dendrites. These data therefore support our conclusion that AP backpropagation into basal dendrites is decremental.

Effect of A-type potassium channels

One possible explanation for the attenuation of single backpropagating APs in distal basal dendrites is that this is attributable to activation of 4-AP-sensitive A-type potassium (I_A) channels, which have been shown previously to limit AP backpropagation in the apical dendrites of both hippocampal and cortical pyramidal neurons (Hoffman et al., 1997; Stuart and Hausser, 2001). To test this, we investigated the impact of 4-AP (4 mM) on AP-evoked dendritic calcium transients detected with the low-affinity calcium-sensitive dye OGB-6F at proximal (<130 μm from the soma) and distal (>130 μm) sites in the basal dendrites. To prevent epileptic discharges in the presence of 4-AP, DNQX (10 μM) and APV (50 μM) were added to the extracellular solution to block AMPA and NMDA receptors, and bicuculline (10 μM) was added to block $GABA_A$ receptors. Single APs were evoked by a depolarizing/hyperpolarizing current sequence to block AP burst firing typically observed after bath application of 4-AP. At proximal sites, the calcium influx associated with backpropagation of single APs, as well as high-frequency AP bursts, was not changed after blockade of I_A channels (Fig. 4*A*). In contrast, bath application of 4-AP significantly enhanced the response to both single APs and AP bursts at distal basal locations ($p < 0.05$) (Fig. 4*B*). To investigate the role of D-type potassium channels, which would be expected to be blocked by high concentrations of 4-AP, we tested the effect of low concentrations of 4-AP (100 μM). To increase the resolution of single AP responses, these experiments were performed using OGB-1. Bath application of low concentrations of 4-AP (100 μM) had no significant effect on the single AP response at distal locations in basal dendrites (>130 μm from the soma; $n = 4$) (Fig. 4*C*). Together, these findings suggest that A-type, fast-inactivating, potassium channels in basal dendrites play an important role in limiting AP backpropagation, as has been observed previously in apical and oblique dendrites of pyramidal neurons (Hoffman et al., 1997; Stuart and Hausser, 2001; Frick et al., 2003).

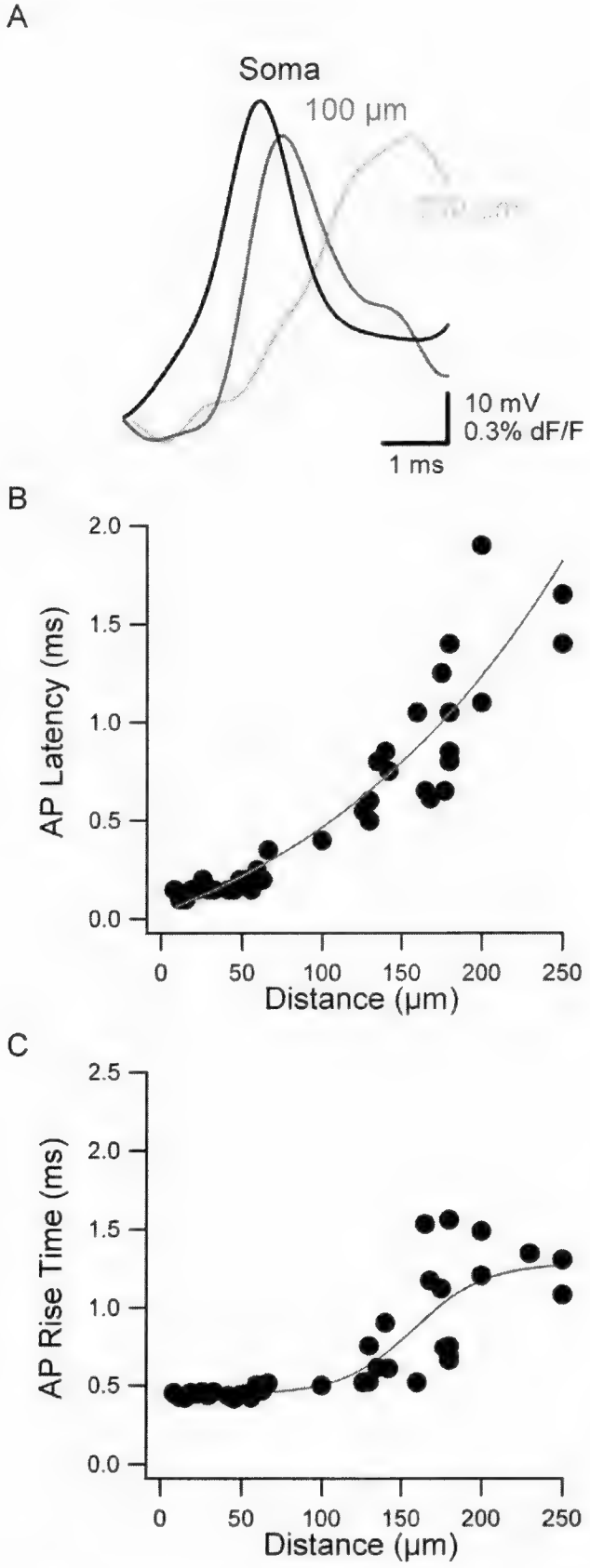


Figure 3. Properties of single backpropagating APs in basal dendrites. ***A***, Superimposed somatic voltage recording (black) and dendritic voltage imaging (gray traces). ***B***, AP peak latency increases in a nonlinear way with distance. Line shows exponential fit. ***C***, Dendritic AP rise time decreases with distance. Line shows sigmoidal fit.

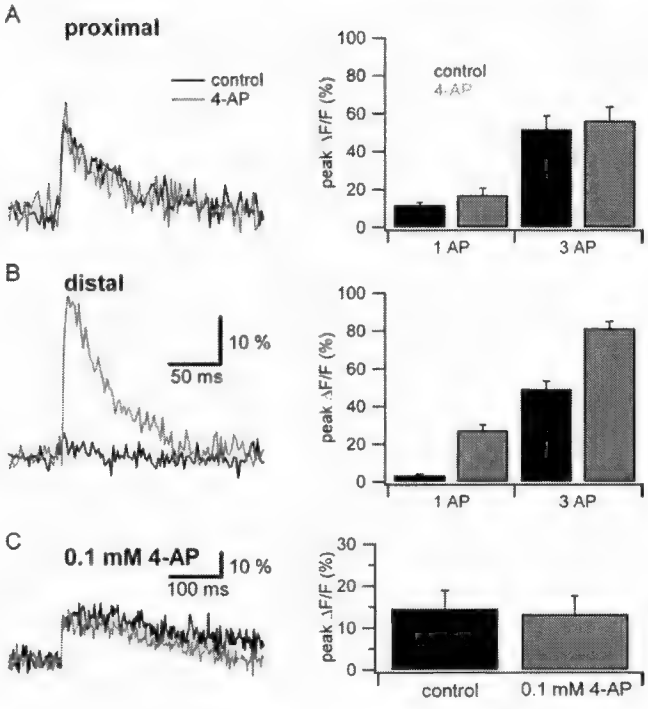


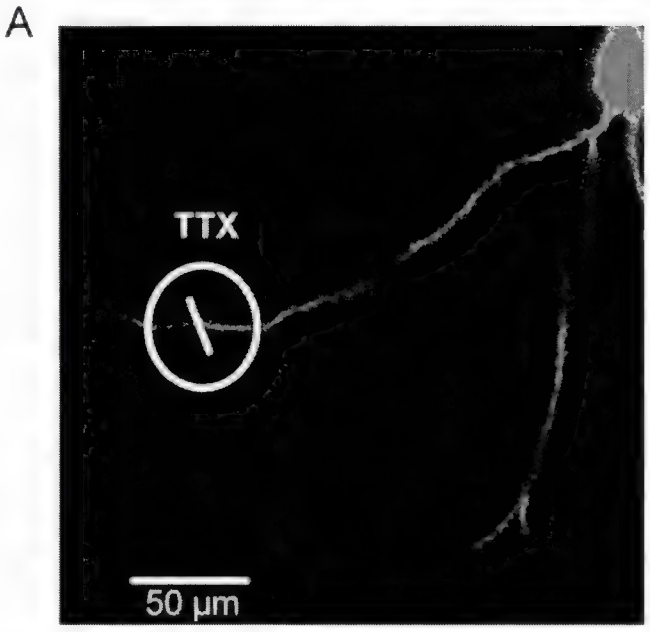
Figure 4. AP backpropagation in distal basal dendrites is modulated by A-type potassium channels. ***A***, Calcium transients evoked by single backpropagating APs in control (gray) and after bath application of 4-AP (gray) at a proximal site (50 μm; ***A***) and a distal dendritic site (204 μm; ***B***) of the same cell. Right bar graphs summarize changes in peak calcium transients evoked by one or three APs during 4-AP application (4 mM) for proximal (<130 μm; $n = 6$) and distal (>130 μm; $n = 13$) sites. Changes in intracellular calcium were imaged with OGB-6F. ***C***, Impact of low (100 μM) concentrations of 4-AP. Left, Calcium transients evoked by single backpropagating APs at distal basal locations in control (gray) and after bath application of 100 μM 4-AP (gray). Right, Pooled data of the average change in peak calcium transients evoked by single APs in control (black) and 100 μM 4-AP (gray). Changes in intracellular calcium were imaged with OGB-1.

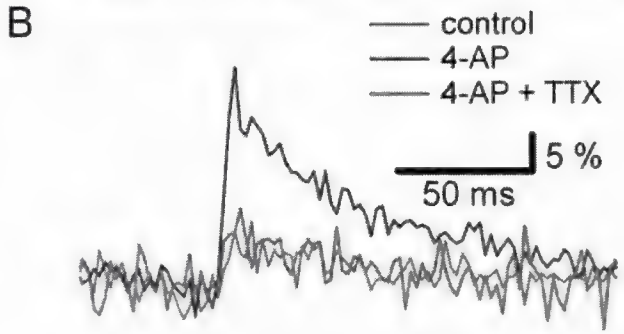
TTX sensitivity of AP backpropagation in distal basal dendrites

To investigate the role of dendritic sodium channels in AP backpropagation, we recorded AP-evoked changes in intracellular calcium during focal applications of TTX (2 μM) in the presence or absence of bath applied 4-AP (Fig. 5*A*). Calcium responses detected with OGB-6F were sequentially imaged in distal basal (>150 μm from the soma) and proximal apical sites during focal application of TTX before, during, and after bath application of 4-AP. Local TTX applications in the absence of 4-AP had no significantly effect on the amplitude of AP-evoked calcium transients at distal basal dendritic sites (Fig. 5*C*) but significantly reduced calcium transients at proximal apical sites (Fig. 5*C*). In contrast, after blockage of I_A channels by bath application of 4-AP (4 mM), focal TTX applications to distal basal sites significantly reduced calcium transients evoked by single APs (Fig. 5*B*,*C*). These results show that, in the presence of 4-AP, backpropagating single APs recruit dendritic sodium channels, whereas this is not the case in the absence of 4-AP. These data suggest that, under control conditions, activation of dendritic I_A channels leads to failure of backpropagation of single APs in distal basal dendrites.

Pairing EPSPs with AP bursts

The results above suggest that AP bursts are likely to provide an effective stimulus for the relief of the voltage-dependent block of synaptic NMDA receptors by magnesium. To investigate this, we





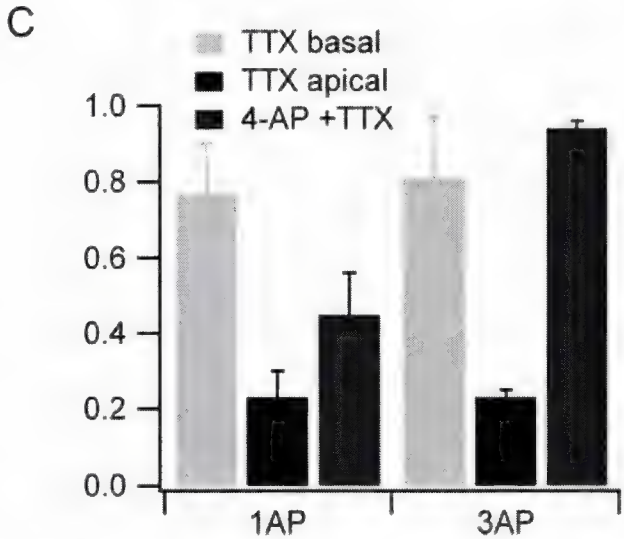


Figure 5. Activation of voltage-gated sodium channels in distal basal dendrites. ***A***, Image of basal dendrite filled with OGB-6F. Line indicates location of confocal line scans, and circle shows location of focal TTX application. ***B***, Calcium transients recorded at a distal basal site after generation of single APs in control (gray), in 4-AP (blue), and during focal application of TTX in the presence of 4-AP (red). ***C***, Summary data showing the amplitude of AP-evoked calcium transients relative to control at distal basal locations during local application of TTX alone (yellow) and in the presence of 4-AP (brown; $n = 6$) during single APs (left) and three AP bursts (right). The amplitude of AP-evoked calcium transients (relative to control) at proximal apical dendritic locations during local TTX applications is shown for comparison (blue).

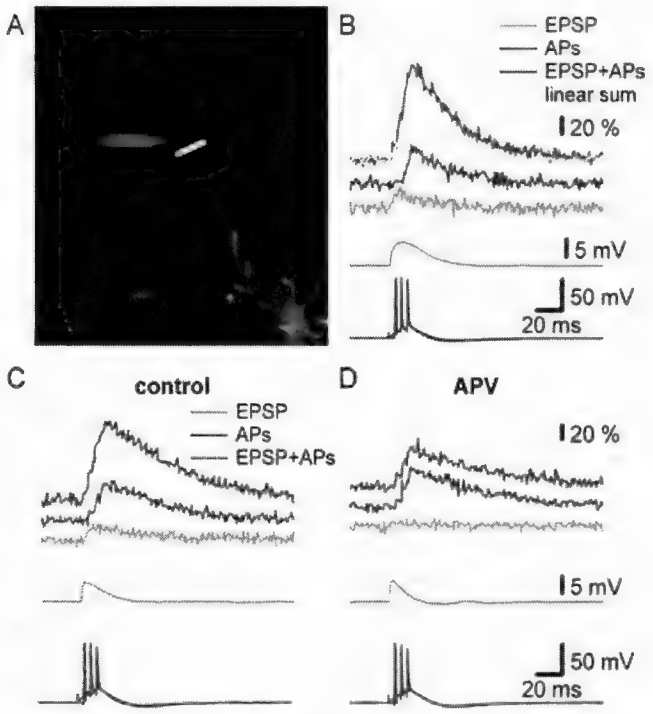


Figure 6. Calcium imaging pairing of AP bursts with EPSPs. ***A***, Image of a layer 5 pyramidal neurons filled with OGB-6F (soma at bottom right). The location of a line scan through a basal dendrite is indicated by the white bar. The extracellular stimulation pipette (red) is positioned in close proximity (<10 μm) to the site in which calcium transients were recorded. ***B***, Dendritic calcium transients ($\Delta F/F$; average of 3 trails) after extracellular stimulation of EPSPs alone (blue), three APs alone (red; 200 Hz), or during pairing of EPSPs and AP bursts at +10 ms (green; middle AP) imaged in a basal dendrite 175 μm from the soma with OGB-6F. Somatic voltage during EPSP/AP pairing is also shown (black). Arrows indicate EPSP onset. Gray traces show the linear sum of the EPSP plus AP response. ***C***, Increase in dendritic calcium ($\Delta F/F$) after stimulation of EPSPs alone (blue), three APs alone (red), or during EPSP/AP pairing at +10 ms (green) in control (top) or in the presence of APV (50 μM; bottom) imaged in a basal dendrite 80 μm from the soma. EPSP response and supralinear increase during pairing are blocked by APV (50 μM).

used calcium imaging to detect changes in intracellular calcium during pairing of EPSP with AP bursts (Fig. 6*A*). These experiments indicated that pairing AP bursts (three APs at 200 Hz) with EPSPs at positive times (+10 ms; middle AP) leads to supralinear increases in intracellular calcium (Fig. 6*B*), which were on average greater at distal dendritic locations (distal, 1.9 ± 0.2-fold, $n = 4$, >150 μm from the soma; proximal, 1.4 ± 0.1-fold, $n = 4$, <150 μm from the soma). Supralinear increases in intracellular calcium during pairing of AP bursts with EPSPs were blocked by the NMDA receptor antagonist D-APV ($n = 3$), indicating the requirement of NMDA receptor activation (Fig. 6*C*). These observations show that appropriately timed AP bursts are effective at activating synaptic NMDA receptors in basal dendrites, suggesting a potential role for AP burst-evoked dendritic calcium spikes in STDP.

Simulations of basal dendritic excitability

Finally, we investigated whether the results above could be simulated using a morphologically realistic model of a layer 5 pyramidal neuron (Fig. 7*A*). Both low- and high-threshold voltage-activated calcium channels were included at a uniform density, because previous experimental studies suggested that they are present in the basal dendrites of layer 5 pyramidal neurons (Schiller et al., 1998). In addition, fast-inactivating A-type potas-

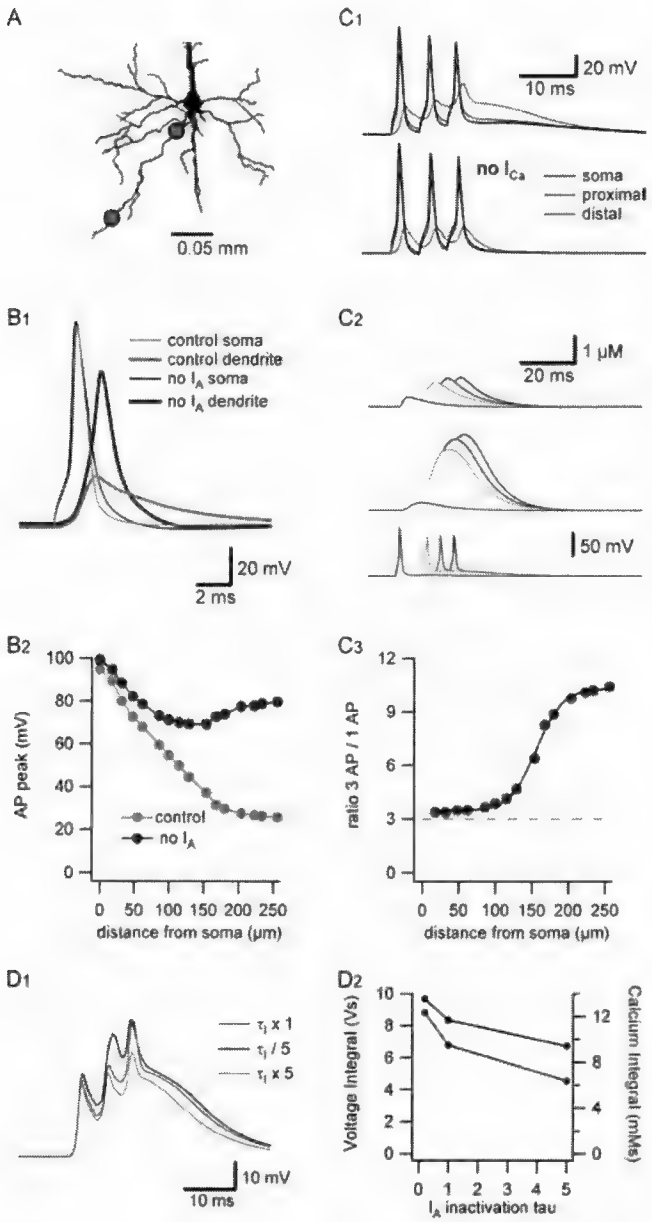


Figure 7. Simulation of AP backpropagation in basal dendrites in a model. ***A***, Image of compartmental model of a layer 5 pyramidal neuron showing the basal dendrites. Circles indicate a proximal (red; 48 μm) and distal (green; 204 μm) dendritic location. ***B***, A-type potassium currents (I_A) reduce AP amplitude in distal basal dendrites. ***B1***, Waveform of single AP at the soma (thin lines) and at a distal dendritic location (204 μm from soma; thick lines) in control (gray traces) and after removal of I_A channels (blue traces). ***B2***, AP amplitude along a basal dendritic branch with (gray symbols) and without (blue symbols) I_A. ***C***, Calcium electrogenesis in basal dendrites. ***C1***, Somatic (black) and dendritic AP waveforms (red, proximal; green, distal, as indicated in ***A***) during a three AP burst (200 Hz) under control conditions (top) and after removal of T-type and high-voltage-activated calcium channels from basal dendrites (bottom). Compare with Figure 2, *D* and *E*. ***C2***, Intracellular calcium concentration at the basal dendritic locations indicated in ***A*** during trains of one to five APs (200 Hz; color-coded traces). Somatic voltage is shown below. Compare with Figure 1*D*. ***C3***, Ratio of the peak calcium transient evoked by three APs relative to one AP at different distances from the soma. Compare with Figure 1*E*. Black line shows sigmoidal fit. Dashed line indicates a linear ratio. ***D1***, Dendritic AP waveforms during three AP bursts (200 Hz) in models with different inactivation time constants of I_A. Recorded 204 μm from the soma. ***D2***, Plot of voltage and calcium integral during AP bursts versus I_A inactivation time constant (relative to control) for data like that in ***D1***.

sium channels were included at a uniform density throughout the basal dendrites. Single APs in this model failed to fully invade the distal region of basal dendrites, whereas removing I_A channels rescued AP backpropagation (Fig. 7*B*). Similar observations were made in models with a linear increase in the density of I_A, as in previous simulations of the role of A-type potassium channels in the regulation of AP backpropagation in hippocampal pyramidal neurons (Migliore et al., 1999). Consistent with our calcium imaging data, these simulations showed that AP bursts generated a dendritic calcium spike that was blocked when dendritic calcium channels were removed (Fig. 7*C1*). In addition, the intracellular calcium concentration at proximal basal sites increased linearly with the number of APs in a burst, whereas at distal basal sites, there was a highly supralinear increase during trains of three or more APs (Fig. 7*C2*). As a consequence, the ratio of the peak calcium transient evoked by three APs relative to that evoked by one AP showed a distance-dependent increase as observed experimentally (compare Figs. 7*C3*, 1*E*). These results were quantitatively similar to the experimental data obtained with calcium (Fig. 1*B–E*) and voltage (Fig. 2*D,E*) imaging.

We next investigated the role of fast inactivation of A-type potassium channels. Models in which the rate of I_A inactivation was slowed by a factor of 5 did not generate dendritic calcium spikes during three AP bursts, whereas speeding up the rate of inactivation by a factor of 5 generated calcium spikes after the second AP in a burst rather than the third (Fig. 7*D1*). These simulations show that the rate of inactivation of A-type potassium channels in basal dendrites plays a critical role in determining the extent of dendritic calcium electrogenesis during AP bursts (Fig. 7*D2*). Together, these findings indicate that the expression of fast-inactivating I_A channels on the one hand and low- and high-voltage-activated calcium channels on the other tune the distal basal dendrites to express dendritic calcium electrogenesis during AP bursts.

Discussion

The majority of synaptic input to cortical pyramidal neurons is made onto their basal dendrites (Larkman, 1991). An understanding of their electrophysiological properties is therefore essential to an understanding of how these neurons process and integrate synaptic input. In addition, the study of AP backpropagation into basal dendrites is particularly relevant to STDP. The experiments presented here demonstrate that single APs are significantly attenuated when invading the distal regions of basal dendrites, whereas AP bursts efficiently depolarize the distal basal dendrites as a result of activation of dendritic calcium channels and the generation of calcium spikes. Differences in the ability of single APs and AP bursts to invade distal basal dendrites were regulated by fast-inactivating, A-type potassium channels and may underlie the importance of AP burst firing for induction of synaptic plasticity (Paulsen and Sejnowski, 2000).

Calcium imaging and electrogenesis in basal dendrites

Previous work in apical dendrites of layer 5 pyramidal neurons indicates that single backpropagating APs evoke only small or negligible increases in intracellular calcium at distal locations, whereas high-frequency bursts lead to substantial distal calcium signals (Larkum et al., 1999). Similarly, in the basal dendrites of layer 5 pyramidal neurons, we found that the calcium signal associated with single backpropagating APs is small at distal sites, whereas the response to AP bursts increases supralinearly with distance of the recording site from the soma (Fig. 1). The observed supralinear increases in calcium are unlikely to be a consequence of

the nonlinear binding properties of the calcium-sensitive dyes used, because similar supralinear increases were observed with high- and low-affinity dyes (Fig. 1*E*). A more likely explanation, which is consistent with data obtained in apical dendrites (Larkum et al., 1999), is that AP bursts cause regenerative activation of dendritic voltage-activated calcium channels, leading to generation of dendritic calcium spikes and substantial calcium influx.

Voltage imaging of AP backpropagation in basal dendrites

Optical methods to measure changes in ion concentrations or membrane voltage offer a great advantage over traditional recording methods when examining events in fine neuronal processes (Tsien, 1989; Denk et al., 1996). Although it is possible to convert fluorescence signals using ion-sensitive dyes into quantitative ion concentrations (assuming the affinity, binding kinetics, and dye concentration are known), it is much more difficult to calibrate signals using VSDs. Dye molecules have to be inserted into the cell membrane to sense the voltage across it. However, when introduced intracellularly, they also bind to the membranes of internal organelles that do not experience voltage changes during AP propagation, contributing to background fluorescence. Furthermore, perfect voltage control of basal dendrites is not possible via somatic recording (Spruston et al., 1993), making quantitative calibration of fluorescence changes at distal dendritic sites impossible. Therefore, only relative changes in dendritic voltage can be measured. In the case of full AP backpropagation, if no attenuation occurs, all APs in a train should have the same amplitude. This was found to be the case in the proximal basal dendrites. In distal basal dendrites, however, the peak of the third AP was much larger than that of the first as a result of generation of an AP burst-evoked dendritic calcium spike (Fig. 2). Given that previous voltage recordings of dendritic calcium spikes during AP bursts indicate that they peak at ~0 mV (Schiller et al., 1997; Larkum et al., 1999), the first AP in the burst is likely to be ~30 mV in amplitude, indicating significant attenuation. This conclusion is consistent with the slower rise time of backpropagating APs at distal basal locations (Fig. 3), the ability of 4-AP to increase the amplitude of distal but not proximal AP-evoked calcium transients (Fig. 4), and the finding that, in absence of 4-AP, local TTX applications did not affect the amplitude of distal dendritic calcium transients during single APs.

The conclusion that single APs attenuate as they propagate into distal basal dendrites is in contrast to that made by Antic (2003). One possible explanation for this difference is that the voltage-sensitive dye imaging in the study by Antic was performed at room temperature rather than physiological temperatures. Alternatively, the discrepancy between our study and that by Antic (2003) may simply be because we imaged more distal dendritic locations. Consistent with this idea, Antic (2003) also showed broadening of backpropagating APs at distal basal dendritic sites in a small subset of cells. It also seems possible that the main difference in the two studies is the conclusion drawn from the data rather than the actual data. No calibration was supplied to determine the amplitude of VSD signals observed in the study of Antic (2003). In our study, however, we were able to compare the relative amplitude of APs in a burst and found that they are not the same, as would be expected if APs propagate without amplitude attenuation.

Mechanisms of AP backpropagation and calcium signals in basal dendrites

In the hippocampus, CA1 neurons express I_A channels at high, non-uniform densities in their apical and oblique dendrites (Hoffman et al., 1997; Frick et al., 2003). Activation of these channels causes a reduction in AP amplitude and the associated calcium signal (Hoffman et al., 1997; Frick et al., 2003). The apical dendrites of layer 5 pyramidal neurons also express I_A channels, albeit at a lower, uniform density (Korngreen and Sakmann, 2000; Bekkers, 2000). Activation of I_A channels contributes to failure of AP backpropagation in the apical dendrites of cortical layer 5 pyramidal neurons (Stuart and Hausser, 2001). Similarly, we show here that the AP-associated calcium influx during single APs in the distal basal dendrites of cortical layer 5 pyramidal neurons is reduced by the activation of I_A channels, because blockade of these potassium channels increased calcium influx triggered by backpropagating APs (Fig. 3). Furthermore, local TTX applications reduced calcium responses during APs when I_A channels were blocked but had no detectable affect under control conditions (Fig. 4). These data suggest that I_A channels contribute to attenuation and possible failure of AP backpropagation in distal basal dendrites.

In contrast, voltage imaging indicated that, at distal dendritic sites, AP bursts produced a large depolarization that was sensitive to blockers of voltage-activated calcium channels. This finding presumably underlies the supralinear calcium influx observed during AP bursts in distal basal dendrites. A likely scenario is that, at distal basal sites, attenuated and broadened backpropagating APs summate with each other when evoked at high frequencies. If the summed voltage reaches threshold for activation of dendritic calcium channels, their activation can become regenerative, leading to generation of a dendritic calcium spike. It has been reported previously that basal dendrites express a range of high-voltage-activated calcium channels, as well as nickel-sensitive low-threshold calcium channels (Schiller et al., 1998). Immunohistochemical studies reveal mainly P/Q-type channels in basal dendrites of layer 5 pyramidal neurons (Westenbroek et al., 1995), whereas N-type channels occur mainly on apical dendrites (Westenbroek et al., 1992) and L-type channels are restricted to the somatic and proximal dendritic region (Hell et al., 1993). Our simulations (Fig. 7) indicate that the expression of both low-threshold T-type calcium channels and high-threshold calcium channels are required for the generation of dendritic calcium spikes in basal dendrites.

Possible role of backpropagating APs in basal dendrites

Backpropagating APs can act as a retrograde messenger informing the input to a neuron, the synapse, about the output of the neuron, the axonal AP. This may play an important role in STDP, in which the relative timing of the postsynaptic AP to the synaptic input determines whether the synapse undergoes potentiation or depression (Bi and Poo, 1998; Sjostrom et al., 2001). AP backpropagation is necessary for the induction of this kind of synaptic plasticity (Magee and Johnston, 1997). If single APs undergo significant attenuation, as appears to be the case in the distal basal dendrites, they will fail to deliver the depolarization needed to release the magnesium block inside the NMDA receptor channel pore. Dendritic calcium spikes evoked by AP bursts, conversely, have been shown to be very effective at removing the voltage-dependent magnesium block of NMDA receptors (Vargas-Caballero and Robinson, 2003) and, as we show here, are effective triggers for NMDA receptor activation (Fig. 6). Together, these findings may explain why AP bursts are necessary for activation of synaptic NMDA receptors in basal dendrites (Kampa et al., 2006) and for the induction of synaptic plasticity in brain slices taken from mature (>3-week-old) rats (Thomas et al., 1998; Pike et al., 1999; Kampa et al., 2006). Conversely, the finding that

single APs at lower frequencies are sufficient for induction of STDP in younger (2-week-old) animals (Markram et al., 1997b; Sjostrom et al., 2001) suggests that backpropagation of single APs into basal dendrites is likely to be more robust at earlier stages of development.

Together, these data suggest that, during STDP, the extent of AP backpropagation and associated dendritic electrogenesis would be expected to play an important role in regulating the degree of synaptic NMDA receptor activation and so plasticity, a mechanism believed to be involved in learning and memory in the brain (Bliss and Collingridge, 1993).

References

Antic SD (2003) Action potentials in basal and oblique dendrites of rat neocortical pyramidal neurons. J Physiol (Lond) 550:35–50.

Bekkers JM (2000) Distribution and activation of voltage-gated potassium channels in cell-attached and outside-out patches from large layer 5 cortical pyramidal neurons of the rat. J Physiol (Lond) 525:611–620.

Bi G-Q, Poo M-M (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci 18:10464–10472.

Bliss TV, Collingridge GL (1993) A synaptic model of memory: long-term potentiation in the hippocampus. Nature 361:31–39.

Denk W, Yuste R, Svoboda K, Tank DW (1996) Imaging calcium dynamics in dendritic spines. Curr Opin Neurobiol 6:372–378.

Frick A, Magee J, Koester HJ, Migliore M, Johnston D (2003) Normalization of Ca^{2+} signals by small oblique dendrites of CA1 pyramidal neurons. J Neurosci 23:3243–3250.

Hell JW, Westenbroek RE, Warner C, Ahlijanian MK, Prystay W, Gilbert MM, Snutch TP, Catterall WA (1993) Identification and differential subcellular localization of the neuronal class C and class D L-type calcium channel alpha 1 subunits. J Cell Biol 123:949–962.

Helmchen F, Imoto K, Sakmann B (1996) Ca^{2+} buffering and action potential-evoked Ca^{2+} signaling in dendrites of pyramidal neurons. Biophys J 70:1069–1081.

Hines ML, Carnevale NT (1997) The NEURON simulation environment. Neural Comput 9:1179–1209.

Hoffman DA, Magee JC, Colbert CM, Johnston D (1997) K^+ channel regulation of signal propagation in dendrites of hippocampal pyramidal neurons. Nature 387:869–875.

Kampa B, Letzkus JJ, Stuart GJ (2006) Requirement of dendritic calcium spikes for induction of spike-timing dependent synaptic plasticity. J Physiol (Lond) 574.1:283–290.

Korngreen A, Sakmann B (2000) Voltage-gated K+ channels in layer 5 neocortical pyramidal neurones from young rats: subtypes and gradients. J Physiol (Lond) 525:621–639.

Larkman AU (1991) Dendritic morphology of pyramidal neurones of the visual cortex of the rat. III. Spine distributions. J Comp Neurol 306:332–343.

Larkum ME, Kaiser KM, Sakmann B (1999) Calcium electrogenesis in distal apical dendrites of layer 5 pyramidal cells at a critical frequency of backpropagating action potentials. Proc Natl Acad Sci USA 96:14600–14604.

Lisman JE (1997) Bursts as a unit of neural information: making unreliable synapses reliable. Trends Neurosci 20:38–43.

Magee JC, Johnston D (1997) A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. Science 275:209–213.

Markram H, Lübke J, Frotscher M, Sakmann B (1997a) Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. Science 275:213–215.

Markram H, Lubke J, Frotscher M, Roth A, Sakmann B (1997b) Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J Physiol (Lond) 500:409–440.

Migliore M, Hoffman DA, Magee JC, Johnston D (1999) Role of an A-type K^+ conductance in the back-propagation of action potentials in the dendrites of hippocampal pyramidal neurons. J Comput Neurosci 7:5–15.

Milojkovic BA, Radojicic MS, Goldman-Rakic PS, Antic SD (2004) Burst generation in rat pyramidal neurones by regenerative potentials elicited in a restricted part of the basilar dendritic tree. J Physiol (Lond) 558:193–211.

Paulsen O, Sejnowski TJ (2000) Natural patterns of activity and long-term synaptic plasticity. Curr Opin Neurobiol 10:172–179.

Pike FG, Meredith RM, Olding AW, Paulsen O (1999) Postsynaptic bursting is essential for "Hebbian" induction of associative long-term potentiation at excitatory synapses in rat hippocampus. J Physiol (Lond) 518:571–576.

Polsky A, Mel BW, Schiller J (2004) Computational subunits in thin dendrites of pyramidal cells. Nat Neurosci 7:621–627.

Sabatini BL, Oertner TG, Svoboda K (2002) The life cycle of Ca^{2+} ions in dendritic spines. Neuron 33:439–452.

Schaefer AT, Larkum ME, Sakmann B, Roth A (2003) Coincidence detection in pyramidal neurons is tuned by their dendritic branching pattern. J Neurophysiol 89:3143–3154.

Schiller J, Schiller Y, Stuart G, Sakmann B (1997) Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. J Physiol (Lond) 505:605–616.

Schiller J, Schiller Y, Clapham DE (1998) NMDA receptors amplify calcium influx into dendritic spines during associative pre- and postsynaptic activation. Nat Neurosci 1:114–118.

Schiller J, Major G, Koester HJ, Schiller Y (2000) NMDA spikes in basal dendrites of cortical pyramidal neurons. Nature 404:285–289.

Sjostrom PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. Neuron 32:1149–1164.

Sjostrom PJ, Turrigiano GG, Nelson SB (2003) Neocortical LTD via coincident activation of presynaptic NMDA and cannabinoid receptors. Neuron 39:641–654.

Snider RK, Kabara JF, Roig BR, Bonds AB (1998) Burst firing and modulation of functional connectivity in cat striate cortex. J Neurophysiol 80:730–744.

Spruston N, Jaffe DB, Williams SH, Johnston D (1993) Voltage- and space-clamp errors associated with the measurement of electrotonically remote synaptic events. J Neurophysiol 70:781–802.

Stuart G, Spruston N (1998) Determinants of voltage attenuation in neocortical pyramidal neuron dendrites. J Neurosci 18:3501–3510.

Stuart G, Spruston N, Sakmann B, Häusser M (1997) Action potential initiation and backpropagation in neurons of the mammalian central nervous system. Trends Neurosci 20:125–131.

Stuart GJ, Hausser M (2001) Dendritic coincidence detection of EPSPs and action potentials. Nat Neurosci 4:63–71.

Stuart GJ, Sakmann B (1994) Active propagation of somatic action potentials into neocortical pyramidal cell dendrites. Nature 367:69–72.

Thomas MJ, Watabe AM, Moody TD, Makhinson M, O'Dell TJ (1998) Postsynaptic complex spike bursting enables the induction of LTP by theta frequency synaptic stimulation. J Neurosci 18:7118–7126.

Tsien RY (1989) Fluorescent probes of cell signaling. Annu Rev Neurosci 12:227–253.

Vargas-Caballero M, Robinson HP (2003) A slow fraction of Mg^{2+} unblock of NMDA receptors limits their contribution to spike generation in cortical pyramidal neurons. J Neurophysiol 89:2778–2783.

Vetter P, Roth A, Hausser M (2001) Propagation of action potentials in dendrites depends on dendritic morphology. J Neurophysiol 85:926–937.

Westenbroek RE, Hell JW, Warner C, Dubel SJ, Snutch TP, Catterall WA (1992) Biochemical properties and subcellular distribution of an N-type calcium channel alpha 1 subunit. Neuron 9:1099–1115.

Westenbroek RE, Sakurai T, Elliott EM, Hell JW, Starr TV, Snutch TP, Catterall WA (1995) Immunochemical identification and subcellular distribution of the α1A subunits of brain calcium channels. J Neurosci 15:6403–6418.

Williams SR, Stuart GJ (2000) Backpropagation of physiological spike trains in neocortical pyramidal neurons: implications for temporal coding in dendrites. J Neurosci 20:8238–8246.

SenseLab ModelDB

Calcium spikes in basal dendrites (Kampa and Stuart 2006)

Accession: 108458

This model was published in Kampa & Stuart (2006) J Neurosci 26(28):7424-32. The simulation creates two plots showing voltage and calcium changes in basal dendrites of layer 5 pyramidal neurons during action potential backpropagation. created by B. Kampa (2006)
Reference: Kampa BM, Stuart GJ (2006) Calcium spikes in basal dendrites of layer 5 pyramidal neurons during action potential bursts. *J Neurosci* **26**:7424-32 [PubMed]

Citations Citation Browser

Model Information (*Click on a link to find other models with that property*)

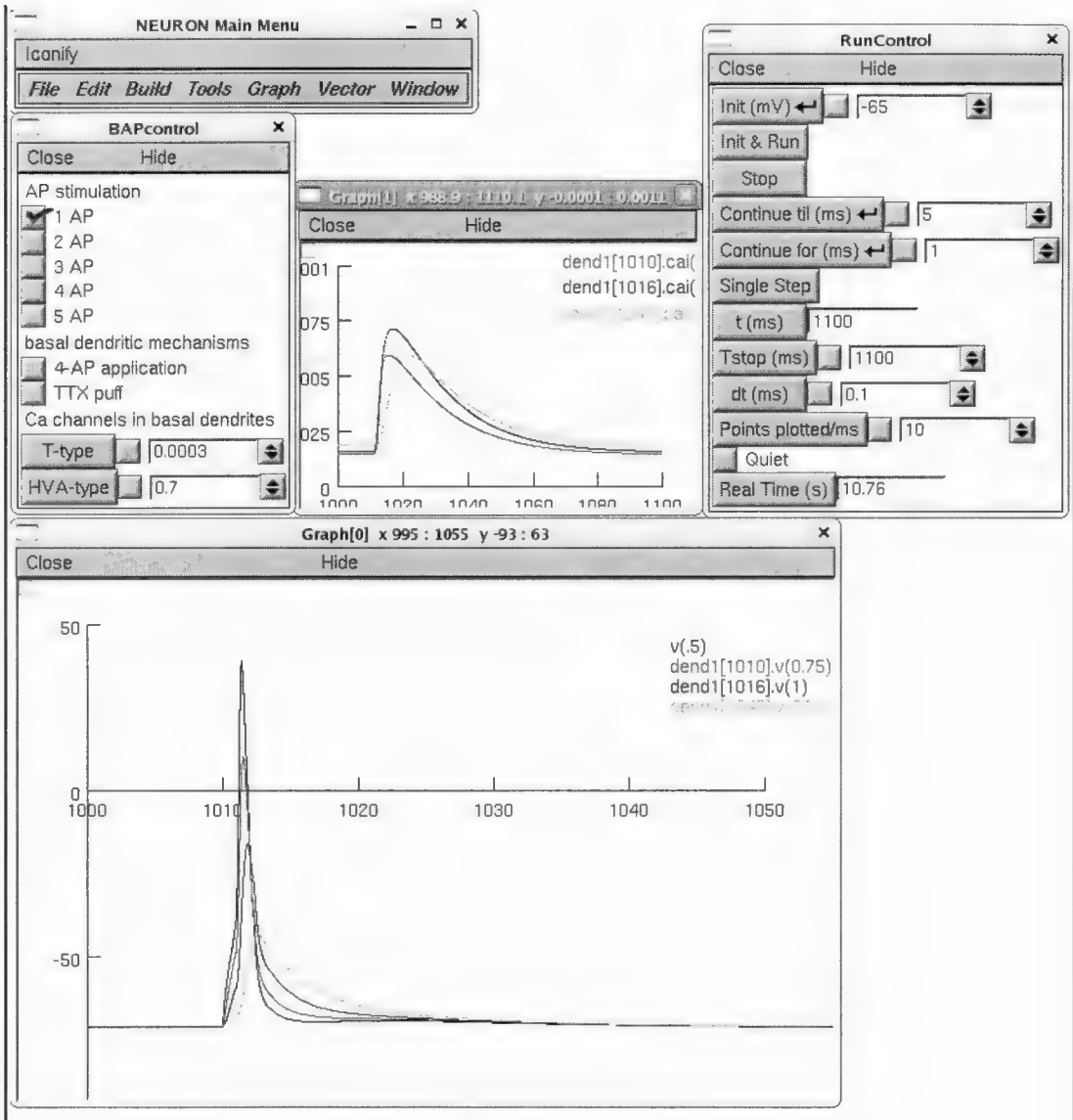
Model Type: Neuron or other electrically excitable cell;
Brain Region(s)/Organism:
Cell Type(s): Neocortical pyramidal neuron: deep;
Channel(s): I L high threshold; I T low threshold; I A; I Sodium; I Potassium;
Gap Junctions:
Receptor(s):
Gene(s):
Transmitter(s):
Simulation Environment: Neuron;
Model Concept(s): Dendritic Action Potentials; Active Dendrites; Action Potentials; Synaptic Integration; Calcium dynamics;
Implementer(s): Kampa, Bjorn M ;

Search NeuronDB for information about: Neocortical pyramidal neuron: deep; I A; I L high threshold; I Potassium; I Sodium; I T low threshold;

Model files Download zip file | Auto-launch **Help downloading and running models**

- \
  - KampaStuart2006
    - mod
    - README.html
    - screenshot.jpg
    - mosinit.hoc
    - BAPcontrol.hoc
    - Ri18geo.hoc
    - Ri18init.hoc
    - Ri18organise.hoc
    - runRi18.hoc
    - BAP_Ca.ses

```
This is the readme for the model published in

Kampa & Stuart (2006) Calcium spikes in basal dendrites of layer 5
pyramidal neurons during action potential bursts.
J Neurosci 26(28):7424-32.

The simulation creates two plots showing voltage and calcium changes
in basal dendrites of layer 5 pyramidal neurons during action
potential (AP) backpropagation.  Running after selecting 1 AP leads
to (after zooming in on the Ca2+ plot):
```

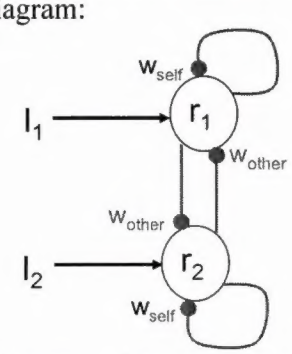



```
Usage:

Autolaunch from ModelDB **or** download and extract the archive and
then compile the mod files in the mod directory and run under ...

linux/unix
----------
by typing
nrnivmodl mod
in the top level directory.  Then type
nrngui mosinit.hoc

mswin
-------

by running mknrndll, cd'ing to the mod directory and press make
nrnmech.dll button.  Copy the nrnmech.dll file to the parent directory
of the mod folder (the top-level directory).  Double click on the
mosinit.hoc file.

MAC OS X
--------

drag and drop the mod folder onto mknrndll icon in the NEURON
application folder.  Start a terminal window, cd to the directory
within the expanded archive and type
mod/powerpc/special mosinit.hoc -
to start the simulation.
```

```
-------

Choose the number of APs in a burst and press the run button. By
increasing the number of APs the calcium signal in the basal dendrites
will increase supralinearly.  3 or more APs in a burst will lead to
generation of a dendritic calcium spike similar to the experimental
results shown in Kampa & Stuart (2006).

In addition, a 4-AP experiment, blocking IA channels, can be simulated
or a TTX experiment blocking local Na channels.

created by B. Kampa (2006)
kampa@hifo.uzh.ch
```

ModelDB Home SenseLab Home Help
Questions, comments, problems? Email the *ModelDB Administrator*
How to cite ModelDB
This site is Copyright 2009 Shepherd Lab, Yale University

1. *Model circuit for selective amplification, neural integration, perceptual rivalry...or many other phenomena.* Many neural networks, including the oculomotor neural integrator, consist of 2 populations of neurons whose effective connectivity is such that they provide self-excitation to neurons in their own population and inhibit neurons in the opposite population. A simple case of these circuits occurs when the connections are symmetric as in the following diagram:



Here, w_{self} gives the connection between either neuron population and itself and w_{other} gives the connection between the different neuron populations. In this problem, we will quite generally find the eigenvalues and eigenvectors for a circuit with this connectivity, and then apply our findings to several relevant cases.

[*Note:* If you have strong familiarity with linear algebra, jump straight to part (d), although try to guess the eigenvectors of **W** rather than calculating them to see if you have developed some network intuition. If you have less background, parts (a)-(c) step you through solving this problem analytically.]

(a) Write down the equations describing this network, assuming the 2 neurons each have intrinsic time constant τ. Write the equations both as separate equations for dr_1/dt and dr_2/dt and as a single equation written compactly in terms of the firing rate vector $\mathbf{r} = [r_1\ r_2]$ and a recurrent connectivity matrix **W**.

(b) Recalling that the eigenvectors of **W** are the patterns of inputs with the property that the outputs of the network will be a scaled version of this same pattern, and noting the symmetry of the network, try to guess the eigenvectors $\xi^{(1)}$ and $\xi^{(2)}$ of **W**. Although MATLAB has a convention of making eigenvectors be of length 1, this is unnecessary and makes the numbers look ugly. Instead, use the convention that the first element of each eigenvector equals +1 [*Note: if you are worried about the generality of this convention, note that any direction except the y-axis itself can be given by a vector (1,y) as long as we realize that a direction with negative x-component like (-1,-1) is really just -1*(1,1), i.e. a vector pointing -1 units along the (1,1) direction.*].

Confirm that you have found the correct direction by showing that $\mathbf{W}\cdot\xi = \lambda\xi$ for each eigenvector and give a formula for the corresponding eigenvalues λ_1 and λ_2.

(c) Decompose the input vector **I** into a combination of the two eigenvectors $\mathbf{I} = b_1\xi^{(1)} + b_2\xi^{(2)}$ you found in part (b). Writing out this equation for each component, you should find that the coefficients b_1 and b_2 respectively correspond to a part of the input that is

common to both cells, I_{common}, and a part I_{diff} that gives the difference of each cell's input from this common value, i.e. that:

$I_1 = I_{common} + I_{diff}$

$I_2 = I_{common} - I_{diff}$

This implies that the common and difference portions of the input behave independently.

(d) *Simple model of selective amplification of differences between inputs.* Suppose that each neuron excites itself ($w_{self} > 0$) and inhibits the other ($w_{other} < 0$) and that the inhibitory connections are stronger than the self-excitatory connections (i.e. $|w_{other}| > w_{self}$).

(i) What will happen to inputs that are common to the two cells? (will they be amplified or attenuated?). Show this by looking at the eigenvalue for the appropriate eigenvector.

(ii) What will happen to inputs that are opposite for the two cells?

(iii) Calculate the eigenvalues for $w_{self} = 0.2$ and $w_{other} = -0.7$.

(iv) Calculate the amplification factors $1/(1-\lambda)$ for each eigenvector, using the values in part (iii) above.

(v) If each cell in the network has an intrinsic time constant τ=18 ms, what will be the corresponding time constants τ_{eff} for each component? From this answer, do you expect the amplified component to change more or less quickly than the attenuated component?

(e) Consider a network with $w_{self} = 0.2$ and $w_{other} = -0.7$, as in part (d). Suppose I_1=63 Hz and I_2 = 57 Hz.

(i) Write these inputs in the form $\mathbf{I} = b_1\xi^{(1)} + b_2\xi^{(2)}$. What are b_1 and b_2?

(ii) Recall that the steady-state firing rate is obtained by simply scaling each component (b_1 or b_2) of **I** by the corresponding amplification factor ($(1-\lambda_1)^{-1}$ or $(1-\lambda_2)^{-1}$, respectively), and then adding together these components, i.e.

$$\mathbf{r} = \frac{b_1}{1-\lambda_1}\xi^{(1)} + \frac{b_2}{1-\lambda_2}\xi^{(2)}$$

What will the steady-state firing rate vector $\mathbf{r}_\infty = [r_{\infty,1}\ r_{\infty,2}]$ equal? For various initial conditions, sketch the expected trajectory of the firing rates as they approach this steady-state value.

2. (a) Simulate the network from question 1(d-e) by numerically solving the equations (do *not* use your results from question 1 here; the point is to "blindly" simulate these equations and confirm that you get the same answer you found analytically in question 1):

$$\tau \frac{dr_1}{dt} = -(1 - w_{self})r_1 + w_{other} r_2 + I_1$$

$$\tau \frac{dr_2}{dt} = -(1 - w_{self})r_2 + w_{other} r_1 + I_2$$

For initial conditions, set $\mathbf{r}(t=0) = \mathbf{I}$, i.e. start out the network at the firing rate values that would have been obtained in steady-state if there were no recurrent connections.

Plot r_1 and r_2 as a function of time for long enough to see them reach steady state. On the same set of axes, plot the eigenvectors $\xi^{(1)}$ and $\xi^{(2)}$. Check that the steady state values that you get are the same as those you calculated analytically in question 1.

You should find that the network amplified the small differences between the inputs—this amplification of the difference between two inputs is why we say such a network mediates *selective amplification* of differences between the inputs.

(b) Modify the above network so that it integrates a quantity proportional to the difference between I_1 and I_2. What is the condition on w_{self} and w_{other} for this to occur? Also observe what happens when the inputs I_1 and I_2 equal zero. Do you observe persistent neural activity? (For comparison, set I_1 and I_2 equal to zero in part (a)).

For this question, don't worry that firing rates can become negative. You can consider **r** to be the firing rate relative to a background level.

(c) *Winner-take-all networks.* For a more extreme example of the rivalry between inputs observed in part (a), see what happens if w_{other} is a large negative value. For this network, add constraints on your neurons such that firing rates that become negative get thresholded to zero and firing rates that exceed some value get set to a saturated firing rate value r_{max}. Experiment with various initial conditions and input values.

Such a network is called a *winner-take-all network* because the loser becomes zero and the winner achieves the maximum value. This model is commonly used as a simple description of phenomena such as perceptual rivalry.